

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

ESCUELA DE POSGRADO



PONTIFICIA
UNIVERSIDAD
CATÓLICA
DEL PERÚ

TITULO DE LA TESIS

Un enfoque de credibilidad bajo espacios de Hilbert y su estimación mediante modelos lineales mixtos

Tesis para optar el grado de Magíster en Estadística

AUTOR

Raúl Alberto Ruiz Arias

ASESOR

Luis Hilmar Valdivieso Serrano

JURADO

Dr. Cristian Luis Bayes Rodríguez

Dr. Luis Hilmar Valdivieso Serrano

Dra. Mery Elizabeth Doig Camino

LIMA-PERÚ

2012

Dedicatoria

A mi hijo Rodrigo y a mi amor Ana, que son el motor en mi vida.

A mis padres Raúl y Juana.

A mis hermanos: Miriam, Liliana, Elcie, Percy, Silvia, Carlos, Mariela, David.

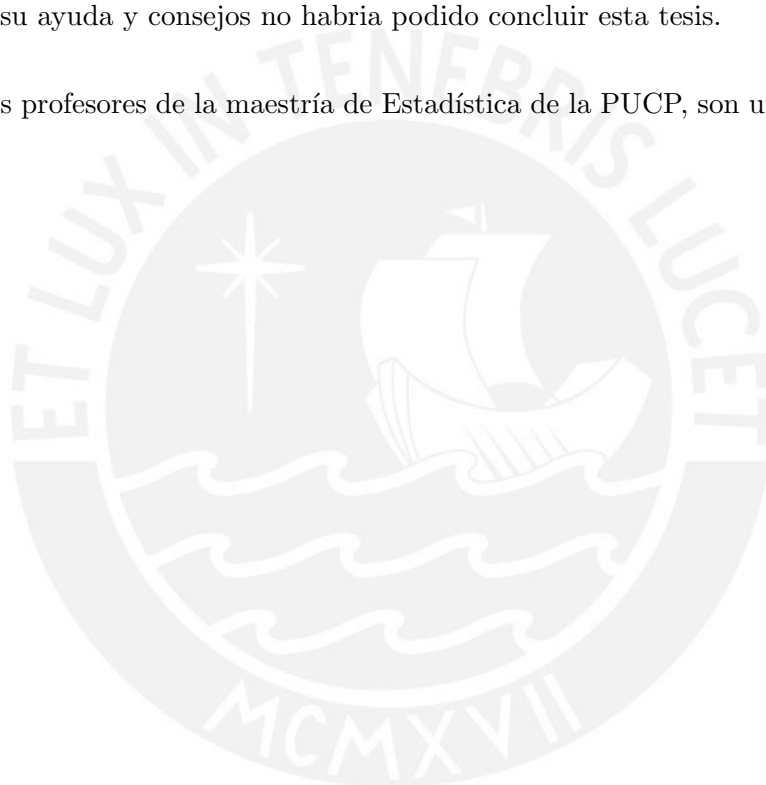


Agradecimientos

A mi papá Raúl que se encuentra en el cielo, gracias por ser un gran ejemplo, gracias por todas tus enseñanzas y por tus sabios consejos.

Muchas gracias a mi asesor de tesis, Dr. Luis Valdivieso, por el gran apoyo que me ha brindado, sin su ayuda y consejos no habría podido concluir esta tesis.

A todos los profesores de la maestría de Estadística de la PUCP, son un gran ejemplo de dedicación.



Índice general

1. Introducción	1
1.1. Consideraciones Preliminares	1
1.2. Objetivos	2
1.3. Organización del Trabajo	2
2. Credibilidad	4
2.1. Preliminares	4
2.2. El modelo de heterogeneidad y el estimador de Bayes	4
2.3. Modelos de credibilidad	9
2.3.1. El modelo de Bühlmann	10
2.3.2. El modelo de Bühlmann-Straub	11
2.3.3. El modelo jerárquico de Jewell	13
2.3.4. El modelo de regresión de Hachemeister	18
2.4. Estimación clásica de los parámetros estructurales	23
2.4.1. Estimadores para el modelo Jerárquico	23
2.4.2. Estimadores del modelo de regresión de Hachemeister	24
2.5. Aplicación con datos de Hachemeister	25
2.5.1. Presentación de los datos	25
2.5.2. Resultados utilizando el modelo de Bühlmann	26
2.5.3. Resultados utilizando el modelo de Bühlmann-Straub	27
2.5.4. Resultados utilizando el modelo de jerárquico de Jewell	28
2.5.5. Resultados utilizando el modelo de regresión de Hachemeister	30
3. Modelos lineales mixtos	32
3.1. El modelo de regresión lineal	32
3.2. El mejor predictor lineal e insesgado de un modelo de regresión	33
3.3. Modelos Lineales Mixtos	35
3.3.1. Estimadores del Modelo	35
4. Credibilidad bajo modelos lineales mixtos	37
4.1. El ajuste al modelo de Bühlmann	37
4.1.1. Aplicación a los datos de Hachemeister	39
4.2. El ajuste al modelo de Bühlmann-Straub	40
4.2.1. Aplicación a los datos de Hachemeister	41
4.3. El ajuste al modelo jerárquico de Jewell	42

4.3.1. Aplicación a los datos de Hachemeister	43
4.4. El modelo de regresión de Hachemeister	44
4.4.1. Aplicación a los datos de Hachemeister	45
5. Estudio de Simulación	46
5.1. Preliminares	46
5.2. Estudio de simulación con el modelo de Bühlmann	46
5.3. Estudio de simulación con el modelo de Bühlmann-Straub	50
5.4. Estudio de simulación con el modelo de Jewell	51
5.5. Estudio de simulación con el modelo de Regresión de Hachemeister	53
6. Conclusiones y Recomendaciones	56
6.1. Conclusiones	56
6.2. Recomendaciones	57
A. Espacios de Hilbert	58
A.1. El espacio L^2	60
A.2. La esperanza condicional	61
A.3. La esperanza condicional con respecto a un conjunto de variables	63
B. Rutinas en R	64
B.1. Programa en R para estimar los parámetros con la data de Hachemeister y el modelo de Bühlmann	64
B.2. Programa en R para estimar los parámetros con la data de Hachemeister y el modelo de Bühlmann-Straub	65
B.3. Programa en R para estimar los parámetros con la data de Hachemeister y el modelo de Jewell	67
B.4. Programa en R para estimar los parámetros con la data de Hachemeister y el modelo de de regresión de Hachemeister	68
B.5. Programa para simular los datos para el modelo de Bühlmann	70
B.6. Programa para simular los datos para el modelo de Buhlmann-Straub	73
B.7. Programa para simular los datos para el modelo de Jewell	75
B.8. Programa en R para simular los datos con el modelo de regresión de Hachemeister	78
Bibliografía	82

Capítulo 1

Introducción

1.1. Consideraciones Preliminares

La idea de poseer un seguro que nos ayude a encarar situaciones inesperadas es bastante antigua, sin embargo la matemática y estadística actuarial tiene en un origen relativamente reciente, es recién a finales de los años 50 del siglo pasado que esta ciencia alcanza propiamente su mayoría de edad.

Uno de los objetivos centrales de esta ciencia es la elaboración de sistemas de tarificación o primas que garanticen la solvencia de una entidad aseguradora pero que simultáneamente resulte lo más justos posible. A principios del siglo XX varios autores propusieron basar esta prima en la historia de reclamaciones y concretamente en la media de estas; sin embargo, se observó empíricamente que con frecuencia este indicador no resultaba ser estable en el tiempo. Es en los años 50 del siglo pasado, que comienza a emerger una corriente estadística que cuestiona algunos de los presupuestos de la escuela ortodoxa, tales como la relevancia de las diversas fuentes de información estadística y la interpretación del concepto de probabilidad. Esta tendencia, que va reinstalando el uso de información inicial en la inferencia estadística, se ha venido asentando con el tiempo y caracteriza a un enfoque de la Estadística, hoy conocido como Escuela Bayesiana. Esta visión se incorporó rápidamente a la ciencia actuarial y fueron Bailey (1945) y Mayerson (1964) los que mostraron como algunas primas obtenidas mediante la metodología Bayesiana podían escribirse como fórmulas de credibilidad; es decir, como una representación de la forma:

$$Prima = (1 - Z)B + Z\bar{X},$$

donde a Z denominaron el factor de credibilidad, el cual ponderaba el conocimiento a priori B con el historial de reclamaciones resumido en la media de reclamaciones \bar{X} .

Esta nueva visión, más amplia, enriqueció la metodología actuarial, proporcionando una fundamentación, si bien parcial, de la que se carecía hasta ese momento en problemas típicamente actuariales como la tarificación a posteriori.

En este trabajo presentaremos algunos de estos modelos de credibilidad, optando por seguir un enfoque geométrico. Fundamentado en la teoría de espacios de Hilbert. Enfocaremos posteriormente la estimación y determinación de primas para estos modelos a través del uso de los modelos lineales mixtos. Mostraremos específicamente que estos modelos pueden escribirse como modelos lineales bajo la inclusión de uno o más efectos aleatorios. Tal caracterización nos permitirá aprovechar toda el potencial y la metodología de los modelos lineales mixtos

(LMM), lo cual resultará, como más adelante veremos en el estudio de simulación, de gran utilidad.

1.2. Objetivos

El objetivo general de la tesis es analizar los principales modelos de credibilidad descritos en la literatura desde un punto de vista geométrico utilizando la teoría de espacios de Hilbert y enfocar la estimación de estos modelos a través del uso de los modelos lineales mixtos (LMM). Más específicamente se busca:

- Introducir, desde un punto de vista geométrico, la teoría de credibilidad y algunos de los principales modelos de credibilidad.
- Revisar la literatura acerca de los modelos lineales mixtos.
- Presentar los modelos de credibilidad como un caso particular de modelos lineales mixtos (LMM).
- Realizar un estudio comparativo de los métodos de inferencia clásica y de inferencia bajo modelos lineales mixtos en la estimación de los parámetros de los modelos de credibilidad, haciendo uso del software estadístico R.

1.3. Organización del Trabajo

En el **primer capítulo**, presentamos una contextualización sobre esta investigación, los objetivos y la estructura de la misma.

En el **segundo capítulo** estudiaremos los elementos necesarios para la comprensión de lo que es un modelo de credibilidad y su posterior enfoque como modelo lineal mixto (LMM). Daremos primero la noción de credibilidad presentando el modelo básico de heterogeneidad y las principales extensiones que abordaremos en este trabajo: el modelo de [Bühlmann \(1967\)](#), [Bühlmann y Straub \(1970\)](#), [Jewell \(1975\)](#) y [Hachemeister \(1975\)](#). Veremos luego como obtener, con la ayuda del teorema de proyección en espacios de Hilbert, las primas de riesgo lineales de Bayes bajo estos modelos y posteriormente como estimar los parámetros estructurales de estas primas en forma clásica. Estas estimaciones se ilustrarán con una base de datos de reclamos de accidentes automovilísticos propuesta por [Hachemeister \(1975\)](#).

El **tercer capítulo** está dedicado al estudio de los modelos lineales mixtos. Empezaremos desarrollando las características de un modelo de regresión lineal múltiple y definiremos luego el modelo lineal mixto mediante la inclusión de los denominados efectos aleatorios. El resto del capítulo estará dedicado al estudio de las características de estos modelos con énfasis en la obtención sus estimadores y valores pronosticados.

En el **cuarto capítulo** retomaremos los modelos de credibilidad introducidos en el segundo capítulo, pero presentados esta vez desde el enfoque de los modelos lineales mixtos. Como ilustración implementaremos y aplicaremos bajo este enfoque la estimación de estos modelos con los datos de [Hachemeister \(1975\)](#).

En el **quinto capítulo** haremos un estudio de simulación para comparar las estimaciones de los parámetros estructurales en las primas lineales de Bayes, tanto bajo el enfoque clásico como el de modelos lineales mixtos utilizando los paquetes **lme** y **actuar** de R.

Finalmente, en el **sexto capítulo** resumiremos los resultados obtenidos en este trabajo y sugeriremos algunas investigaciones futuras que se puedan derivar de esta investigación.

En el apéndice **A** presentamos la teoría de espacios de Hilbert y en particular el teorema de proyección en estos espacios, el cual es utilizado recurrentemente en este trabajo. Incluimos también un apéndice **B** en donde detallaremos las distintas rutinas desarrolladas en lenguaje R.



Capítulo 2

Credibilidad

2.1. Preliminares

La teoría de la credibilidad es una rama de las matemáticas actuariales que construye y analiza modelos que nos permiten la selección de una prima adecuada para un portafolio de clientes. Estas primas se obtienen tomando en cuenta tanto la información histórica (de montos o cantidades de reclamaciones del portafolio) así como la información proveniente de otras fuentes.

Para entender los modelos de credibilidad que desarrollaremos en este trabajo, y que luego enfocaremos en su estimación desde los modelos lineales mixtos, hemos dedicado un apéndice A que introduce el concepto de espacios de Hilbert y otros conceptos relacionados como los espacios L^2 y la esperanza condicional. De vital importancia en este desarrollo lo constituye el teorema de proyección para espacios de Hilbert. Este, que es una extensión natural en estos espacios de la clásica proyección de un vector sobre un plano, será una herramienta ampliamente utilizada a lo largo de este trabajo.

2.2. El modelo de heterogeneidad y el estimador de Bayes

Estaremos interesados en estudiar el comportamiento del promedio de un portafolio de seguros que pudiera involucrar a varias pólizas o contratos a las cuales denotaremos con el subíndice i . Valga aclarar aquí, que por póliza entenderemos no necesariamente a un afiliado individual sino en general a un grupo de afiliados que tienen un vínculo común y que están por tanto afectos a una prima específica. Asumiremos también de que disponemos de data histórica de las severidades (montos o cantidades de reclamos) de estas pólizas a lo largo de un cierto número de periodos, los cuales podrían medirse en meses, semestres, años, etc. Estos se representarán a través de la secuencia de variables aleatorias:

$$X_{i1}, X_{i2}, \dots, X_{in_i},$$

donde X_{it} se interpreta como la severidad de la póliza i en el periodo t y n_i denota al número de observaciones con que contamos del historial de estos reclamos.

Con el fin de obtener una prima base para la póliza i podríamos optar por considerar, como se hace clásicamente, al monto esperado de las severidades de esta póliza. Si asumimos

independencia, tendremos entonces por la ley fuerte de los grandes números, que

$$\bar{X}_i = \frac{1}{n_i} \sum_{t=1}^{n_i} X_{it} \xrightarrow{c.p.} E[X_{it}]$$

podría servirnos de aproximación para la prima base. Esto requiere que n_i sea grande y por tanto que dispongamos de información sobre la póliza i .

Si n_i es pequeño la aproximación anterior podría resultar poco precisa y por tanto el actuario debería pensar en utilizar adicionalmente y/o alternativamente otro tipo de información. Bühlmann (1967) propuso para tal fin un "parámetro" Θ_i asociado a cada póliza i de tal manera que esta contenga información esencial de ella. Dado que por lo usual uno desconoce las características de esta información antes que la póliza se adquiera, se asume que la secuencia (Θ_i) , donde Θ_i corresponde a la póliza i , constituye una secuencia de variables aleatorias independientes e idénticamente distribuidas (i.i.d). Esto significa que todas las pólizas se comportan en promedio de la misma manera; lo que importa es la realización $\Theta_i(\omega)$ de la variable aleatoria Θ_i que determina las propiedades individuales de la póliza i , y la totalidad de los valores de $\Theta_i(\omega)$ determina la heterogeneidad del portafolio. De manera más formal, el modelo de heterogeneidad, como lo plantea Mikosch (2003) y también se describe en Valdivieso (2011), asume que cada póliza i puede ser descrita por el par

$$(\Theta_i, (X_{it})_{t \geq 1}),$$

donde Θ_i es una v.a. que la llamaremos la variable estructural de la póliza i y

$$(X_{it})_{t \geq 1} = (X_{i1}, X_{i2}, X_{i3}, \dots)$$

es una secuencia de severidades históricas (cantidades o montos de reclamos) de la póliza i . Adicionalmente el modelo de heterogeneidad asume los siguientes supuestos.

Definición 2.1 (El modelo de heterogeneidad) .

- (H_1) Los pares $(\Theta_i, (X_{it})_{t \geq 1})$, para $i=1, 2, 3, \dots$, son mutuamente independientes e idénticamente distribuidos (i.i.d).
- (H_2) Dado Θ_i , la secuencia de variables aleatorias $(X_{it})_{t \geq 1}$ es i.i.d con función de distribución común $F(\cdot | \Theta_i)$.

Las condiciones en este modelo implican que el historial de reclamos de la póliza, esta dado por la secuencia de severidades, las que son mutuamente independiente de las demás pólizas. Esta es una condición natural que dice que las diferentes pólizas no interfieren entre sí. La dependencia es solo posible entre las cantidades reclamadas X_{it} al interior de cada póliza i ; sin embargo vale aclarar aquí que estas mismas severidades resultan ser idénticamente distribuidas desde que para cualquier t :

$$P(X_{it} \leq x) = \int P(X_{it} \leq x | \Theta_i = u) dF(u | \Theta_i) = \int P(X_{i1} \leq x | \Theta_i = u) dF(u | \Theta_i) = P(X_{i1} \leq x).$$

La suposición (H_2) de que estas variables aleatorias son i.i.d dado Θ_i es sin duda una idealización que se ha hecho por conveniencia matemática. En adelante llamaremos a los parámetros (no aleatorios) de la variable estructural Θ_i los parámetros estructurales.

En la teoría clásica $E[X_{it}]$ es una medida de riesgo común para todo el portafolio. Esta no toma en cuenta la heterogeneidad del portafolio y la información particular de cada póliza. Una medida de riesgo más adecuada para una póliza particular i vendrá dada por:

$$\begin{aligned}\mu(\Theta_i) &= E[X_{i1}|\Theta_i] \\ &= \int_{\mathbb{R}} x dF(x|\Theta_i).\end{aligned}$$

Ella se denomina la prima de riesgo individual de la póliza i , la cual asumiremos que está bien definida, para lo cual requeriremos que $E[X_{i1}] < \infty$. Dado que $\mu(\Theta_i)$ es una función medible de la variable aleatoria Θ_i y la secuencia (Θ_i) es i.i.d, las secuencias $(\mu(\Theta_i))$ y $(\text{Var}(X_{i1}|\Theta_i))$, última que denotaremos de aquí en adelante por $(\sigma^2(\Theta_i))$, resultan ser también i.i.d.

Un objetivo central en la teoría de credibilidad es encontrar una aproximación razonable a la prima de riesgo individual $\mu(\Theta_i) = E[X_{i1}|\Theta_i]$ en base a las observaciones de los vectores o secuencias de severidades históricas, a denotarse de aquí en adelante por

$$\mathbf{X}_i = (X_{it})_{t \geq 1},$$

de las k pólizas que supondremos conforman un portafolio. Dado que por (H_1) estos vectores son independientes, parece poco probable que \mathbf{X}_j , con $j \neq i$ contenga información útil acerca de $\mu(\Theta_i)$. Más adelante confirmaremos esta sospecha.

En lo que sigue, supondremos que $\text{Var}(\mu(\Theta_i)) < \infty$, por lo que tiene sentido considerar la función

$$\rho(\mu) = E[(\mu(\Theta_i) - \mu)^2], \quad (2.1)$$

donde $\mu = g(\mathbf{X}_1, \dots, \mathbf{X}_k)$ es una función medible de valor real con varianza finita de los vectores $\mathbf{X}_1, \dots, \mathbf{X}_k$. Vale aclarar que la notación $\rho(\mu)$ puede parecer un poco engañosa ya que ρ no es sólo una función de la variable aleatoria μ , sino del vector $(\mu, \mu(\Theta_i))$. No obstante, a menos que sea necesario hacerlo explícito, nosotros seguiremos utilizando por simplicidad esta notación. Llamaremos a la cantidad $\rho(\mu)$ el riesgo o error cuadrático medio de μ (con respecto a $\mu(\Theta_i)$).

Teorema 2.1 (Estimador de Bayes) *El mínimo riesgo sobre la clase de todas las funciones medibles con varianza finita de $\mathbf{X}_1, \dots, \mathbf{X}_k$ existe y es único con probabilidad 1. Este es obtenido tomando en (2.1):*

$$\hat{\mu}_i^B = E[\mu(\Theta_i)|\mathbf{X}_i]$$

y se cumple que:

$$\rho(\hat{\mu}_i^B) = E[\text{Var}(\mu(\Theta_i)|\mathbf{X}_i)] = \text{Var}(\mu(\Theta_i)) - \text{Var}(\hat{\mu}_i^B).$$

Llamaremos a $\hat{\mu}_i^B$ el estimador de Bayes de la prima de riesgo individual $\mu(\Theta_i)$.

Demostración: Puesto que $\mu(\Theta_i) \in L^2(\Omega, \mathcal{F}, P)$ (véase el apéndice A para este y otros conceptos), podríamos considerar la sub- σ -álgebra \mathcal{G} de \mathcal{F} generada por $\mathbf{X}_1, \dots, \mathbf{X}_k$ y el subespacio \mathbf{G} de L^2 de todas las variables aleatorias \mathcal{G} -medibles. Dado que el límite de variables \mathcal{G} -medibles es también \mathcal{G} -medible, \mathbf{G} es un subespacio cerrado de L^2 . En consecuencia, podemos aplicar el teorema de proyección en espacios de Hilbert para garantizar que con probabilidad 1, $E[\mu(\Theta_i)|\mathbf{X}_1, \dots, \mathbf{X}_k]$ es la única variable aleatoria en \mathbf{G} que minimiza la función de riesgo (2.1). Más aún, como por (H_1) Θ_i es independiente de \mathbf{X}_j , $\forall j \neq i$, se tiene que $E[\mu(\Theta_i)|\mathbf{X}_1, \dots, \mathbf{X}_k]$ no es sino el estimador de Bayes $\hat{\mu}_i^B$ dado en el teorema. De otro lado, el riesgo bajo el estimador de Bayes viene dado por:

$$\begin{aligned} \rho(\hat{\mu}_i^B) &= E[(\mu(\Theta_i) - \hat{\mu}_i^B)^2] \\ &= E[E[(\mu(\Theta_i) - E[\mu(\Theta_i)|\mathbf{X}_i])^2]|\mathbf{X}_i] \\ &= E[\text{Var}(\mu(\Theta_i)|\mathbf{X}_i)] \\ &= E[E[\mu(\Theta_i)^2|\mathbf{X}_i] - (\hat{\mu}_i^B)^2] \\ &= E[\mu(\Theta_i)^2] - E[(\hat{\mu}_i^B)^2] \\ &= \text{Var}(\mu(\Theta_i)) - \text{Var}(\hat{\mu}_i^B) \end{aligned}$$

□

El ejemplo siguiente muestra explícitamente el cálculo de un estimador de Bayes bajo algunas asunciones sobre las distribución del número de reclamos (o montos de valor 1) y de la variable estructural.

Ejemplo Consideremos aquí un portafolio con tan sólo una póliza, por lo que para simplificar omitiremos el subíndice i . Supongamos que los números de reclamos $X_t, t = 1, 2, \dots, n$ son v.a.'s i.i.d con distribución de Poisson dado Θ y $\Theta \sim \Gamma(\alpha, \beta)$ para algunos valores positivos α y β . El estimador de Bayes $\hat{\mu}^B = E[\Theta|\mathbf{X}]$ de $\mu(\Theta) = E[X_1|\Theta] = \Theta$ viene dado por el valor esperado de la distribución de $\Theta|X$, donde $\mathbf{X} = (X_1, X_2, \dots, X_n)$ representa al número de reclamos históricos. Esta distribución para las severidades observadas $\mathbf{x} = (x_1, x_2, \dots, x_n)$ satisface, por el teorema de Bayes:

$$\begin{aligned} f_\Theta(\theta|\mathbf{X} = \mathbf{x}) &= P(\mathbf{X} = \mathbf{x}|\theta = \mathbf{x})f_\Theta(\theta) \\ &= P(X_1 = x_1|\Theta = \theta) \dots P(X_n = x_n|\Theta = \theta)f_\Theta(\theta) \\ &\propto \theta^{\alpha-1} e^{-\beta\theta} \prod_{t=1}^n \left(\frac{\theta^{x_t}}{x_t!} e^{-\theta} \right) \\ &\propto \theta^{\alpha + \sum_{t=1}^n x_t - 1} e^{-\theta(\beta + n)} \end{aligned}$$

donde \propto indica que la cantidad a izquierda es proporcional al de la derecha o igual si lo

multiplicamos por un factor que no dependa de θ . Puesto la última expresión representa una función de densidad es posible concluir, a partir de su forma particular, que $\Theta|\mathbf{X} = \mathbf{x}$ tiene una distribución $\Gamma(\alpha + \sum_{t=1}^n x_t, \beta + n)$. Así el estimador de Bayes $\hat{\mu}^B = E[\Theta|\mathbf{X}]$ viene dado explícitamente por:

$$\hat{\mu}^B = \frac{\alpha + \sum_{t=1}^n X_t}{\beta + n}$$

y su riesgo correspondiente es

$$\begin{aligned} \rho(\hat{\mu}^B) &= E[Var(\theta|X)] \\ &= E\left[\frac{\alpha + \sum_{t=1}^n X_t}{(\beta + n)^2}\right] \\ &= \frac{\alpha + nE[E[X_1|\Theta]]}{(\beta + n)^2} \\ &= \frac{\alpha + nE[\Theta]}{(\beta + n)^2} \\ &= \frac{\alpha}{\beta} \left(\frac{1}{\beta + n}\right). \end{aligned}$$

Resulta revelador apreciar que el estimador de Bayes puede también escribirse como:

$$\hat{\mu}^B = (1 - Z)E[\Theta] + Z\bar{X},$$

donde

$$\bar{X} = \frac{1}{n} \sum_{t=1}^n X_t \quad \text{y} \quad Z = \frac{n}{\beta + n},$$

término último que es conocido como el factor de credibilidad. Así, el estimador de Bayes es una media ponderada de la media muestral de la póliza individual y la esperanza del parámetro de heterogeneidad $E[\Theta]$. Note que $Z \rightarrow 1$ si el tamaño de la muestra $n \rightarrow \infty$. Esto significa que el estimador de Bayes $\hat{\mu}_B$ es más cercano a \bar{X} cuanto mayor sea el tamaño de la muestra. Para n pequeños, la variación de \bar{X} es demasiado grande para ser representativa de la póliza. Por lo tanto el peso Z dado para la media de las severidades \bar{X} es pequeño, mientras que el peso $1 - Z$ asignado al valor esperado $E[\Theta]$ es cercano a uno. Esto significa que la prima neta representada por $\mu(\Theta) = E[X_1|\Theta] = \Theta$ está fuertemente influenciada por la información disponible en la póliza. En particular, si no es disponible dicha información; es decir, $n = 0$, el cálculo de la prima se basa únicamente en la expectativa general del portafolio. Observe también que el riesgo satisface:

$$\begin{aligned} \rho(\hat{\mu}^B) &= (1 - Z)Var(\Theta) \\ &= (1 - Z)\frac{\alpha}{\beta^2} \rightarrow 0, \text{ cuando } n \rightarrow \infty. \end{aligned}$$

2.3. Modelos de credibilidad

Un problema con el estimador de Bayes es que la minimización sobre la clase de todas las funciones medibles de la data de severidades con varianza finita pocas veces nos lleva a encontrar de manera explícita a $\hat{\mu}_B$. El ejemplo previo es uno de estos casos especiales y en él el estimador de Bayes resulta ser lineal en el sentido de que es una combinación afin-lineal de las severidades disponibles. Con el fin de obtener un estimador de Bayes analítico muchos autores han propuesto restringir la minimización de la función de riesgo (2.1) sobre una clase menor. Una clase potencialmente útil para tal fin, como lo ha demostrado el ejemplo anterior, es la clase de estimadores lineales:

$$\{c_{i0} + \sum_{t=1}^k \sum_{t=1}^{n_i} c_{it} X_{it} / c_{i0}, c_{it} \in \mathbb{R}, \forall i = 1, 2, \dots, k \text{ y } t = 1, \dots, n_i\}.$$

Los modelos que seguidamente describiremos consideran todos, bajo distintas suposiciones, a estimadores de Bayes restringidos a esta clase. A estos los denominaremos estimadores lineales de Bayes y los denotaremos, en el caso de una póliza particular i , por $\hat{\mu}_i^{LB}$. Note además, por el teorema 2.1, que la minimización de la función de riesgo (2.1) para esta póliza bastará realizarse sobre la clase más restringida

$$\mathcal{L} = \{c_{i0} + \sum_{t=1}^{n_i} c_{it} X_{it} / c_{i0}, c_{it} \in \mathbb{R} \forall t = 1, \dots, n_i\}.$$

Por construcción, un estimador lineal de Bayes, $\hat{\mu}_i^{LB}$, es el mejor estimador (en el sentido medio cuadrático) en \mathcal{L} de la prima de riesgo individual $\mu(\Theta_i) \in L^2(\Omega, \mathcal{F}, P)$. Así el teorema de proyección en espacios de Hilbert nos garantiza que

$$E[(\mu(\Theta_i) - \hat{\mu}_i^{LB})Y] = 0, \forall Y \in \mathcal{L}.$$

Dado que $\hat{\mu}_i^{LB}$ toma la forma $\hat{\mu}_i^{LB} = \hat{c}_{i0} + \sum_{t=1}^{n_i} \hat{c}_{it} X_{it}$, para ciertas constantes $\hat{c}_{i0}, \hat{c}_{i1}, \dots, \hat{c}_{in_i}$ y toda constante no nula c así como las variables X_{is} pertenecen a \mathcal{L} se debe cumplir que

$$E[(\mu(\Theta_i) - \hat{\mu}_i^{LB})c] = 0$$

y

$$E[(\mu(\Theta_i) - \hat{\mu}_i^{LB})X_{is}] = 0, \forall s = 1, 2, \dots, n_i.$$

Esto nos conduce a las siguientes relaciones generales de gran utilidad para obtener los estimadores lineales de Bayes:

$$\hat{c}_{i0} = m(1 - \sum_{t=1}^{n_i} \hat{c}_{it}) \quad (2.2)$$

y

$$E[\mu(\Theta_i)X_{is}] = m \hat{c}_{i0} + \sum_{t=1}^{n_i} \hat{c}_{it} E[X_{it}X_{is}], \forall s = 1, 2, \dots, n_i, \quad (2.3)$$

donde $m = E[X_{it}]$.

2.3.1. El modelo de Bühlmann

Bühlmann (1967) adaptó el modelo de heterogeneidad descrito en la sección anterior a los datos de una cartera o portafolio de k contratos o pólizas. La i -ésima póliza en él es, al igual que en el modelo de heterogeneidad, descrita por el vector aleatorio

$$(\Theta_i, \mathbf{X}_i),$$

donde $\mathbf{X}_i = (X_{it})_{t=1,2,\dots,n_i}$; sin embargo, este modelo asume ahora los supuestos siguientes:

(B_1) Las pólizas $i = 1, \dots, k$; es decir, los pares (Θ_i, \mathbf{X}_i) , son independientes y las variables aleatorias Θ_i son idénticamente distribuidas.

(B_2) Dado Θ_i , las v.a.'s \mathbf{X}_i son i.i.d con función de distribución común $F(\cdot|\Theta_i)$.

(B_3) $\forall i = 1, \dots, k$ y $t, s = 1, \dots, n_i$:

$$\text{Cov}(X_{it}, X_{is}|\Theta_i) = \mathbf{1}_{\{t=s\}}\sigma^2(\Theta_i).$$

Puesto que en el modelo de Bühlmann las sucesiones de v.a.'s $(\mu(\Theta_i))$ y $(\sigma^2(\Theta_i))$ son i.i.d., denotaremos en adelante a sus medias y varianzas comunes mediante los siguientes parámetros estructurales:

$$m = E[\mu(\Theta_i)], \quad a = \text{Var}(\mu(\Theta_i)) \quad \text{y} \quad s^2 = E[\sigma^2(\Theta_i)].$$

Aquí m representa una prima de riesgo colectivo (para todo el portafolio), a es un parámetro de heterogeneidad del portafolio y s^2 representa la variación media de siniestralidad al interior de cada póliza.

Note que el modelo de Bühlmann difiere del modelo de heterogeneidad (el cual es un caso particular de él) en lo siguiente:

- Los vectores aleatorios en la secuencia (\mathbf{X}_i) no tienen necesariamente la misma distribución.
- Solamente la esperanza condicional $\mu(\Theta_i)$ y la varianza condicional $\sigma^2(\Theta_i)$ son los mismos para cada X_{it} , con $t = 1, 2, \dots, n_i$. Las otras características distribucionales de los X_{it} no son fijas.

El resultado siguiente, cuya demostración la reservaremos para un modelo más general, nos provee del estimador lineal de Bayes bajo este modelo.

Teorema 2.2 *Bajo los supuestos (B_1) a (B_3), el estimador lineal de Bayes viene dado por la fórmula de credibilidad:*

$$\hat{\mu}_i^{LB} = (1 - Z_i)m + Z_i\bar{X}_i,$$

donde

$$Z_i = \frac{an_i}{s^2 + an_i} \quad \text{y} \quad \bar{X}_i = \frac{1}{n_i} \sum_{t=1}^{n_i} X_{it}.$$

Más aún, el riesgo de Bayes en este modelo viene dado por:

$$\rho(\hat{\mu}_i^{LB}) = (1 - Z_i)a.$$

2.3.2. El modelo de Bühlmann-Straub

El modelo de Bühlmann y Straub (1970) es una extensión del modelo de Bühlmann que permite la introducción de pesos para cada una de las severidades. A las severidades X_{it} les corresponde un peso ω_{it} (ver cuadro 2.1). Estos pesos son usualmente asociados al tamaño o volumen de la póliza. Por ejemplo, uno podría pensar en ω_{it} como el tamaño de una casa que ha sido asegurada contra incendios, el número de individuos que integra una familia asegurada o el tamaño de un auto particular asegurado.

		Parámetro estructural	Periodos						Totales
			1	2	...	t	...	n_i	
Pólizas	1	Θ_1	ω_{11}	ω_{12}	...	ω_{1t}	...	ω_{1n_1}	$\omega_{1.}$
	2	Θ_2	ω_{21}	ω_{22}	...	ω_{2t}	...	ω_{2n_2}	$\omega_{2.}$
	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots	\vdots
	i	Θ_i	ω_{i1}	ω_{i2}	...	ω_{it}	...	ω_{in_i}	$\omega_{i.}$
	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots	\vdots
	k	Θ_k	ω_{k1}	ω_{k2}	...	ω_{kt}	...	ω_{kn_k}	$\omega_{k.}$

Cuadro 2.1: Pesos para un portafolio de k pólizas en donde $\omega_{i.} = \sum_{t=1}^{n_i} \omega_{it}$.

El modelo de Bühlmann-Straub asume básicamente asunciones similares a las vistas en el modelo de Bühlmann, el cual es un caso particular de este de asumirse pesos iguales a 1; sin embargo a diferencia del segundo la varianza de las severidades conocido el parámetro de heterogeneidad varía en este modelo según los pesos conocidos. Concretamente el modelo de Bühlmann-Straub se basa en los siguientes supuestos:

(BS₁) Las pólizas $i = 1, \dots, k$; es decir, los pares (Θ_i, \mathbf{X}_i) , son independientes y las variables aleatorias Θ_i son idénticamente distribuidas.

(BS₂) Dado Θ_i , la secuencia de variables aleatorias en \mathbf{X}_i , son independiente.

(BS₃) $\forall i = 1, \dots, k$ y $t, s = 1, \dots, n_i$:

$$Cov(X_{it}, X_{is} | \Theta_i) = \mathbf{1}_{\{t=s\}} \frac{1}{\omega_{it}} \sigma^2(\Theta_i).$$

Bajo las mismas notaciones para los parámetros estructurales que en el modelo de Bühlmann, tenemos el siguiente resultado.

Teorema 2.3 *Bajo los supuestos (BS₁) a (BS₃), el estimador lineal de Bayes viene dado por la fórmula de credibilidad:*

$$\hat{\mu}_i^{LB} = (1 - Z_i)m + Z_i \bar{X}_{i\omega},$$

donde

$$Z_i = \frac{a\omega_{i.}}{a\omega_{i.} + s^2} \quad y \quad \bar{X}_{i\omega} = \frac{1}{\omega_{i.}} \sum_{t=1}^{n_i} \omega_{it} X_{it}.$$

Más aún, el riesgo de Bayes en este modelo viene dado por:

$$\rho(\hat{\mu}_i^{LB}) = (1 - Z_i)a.$$

Demostración: Como vimos en la sección 2.3, el estimador lineal de Bayes tiene la forma $\hat{\mu}_i^{LB} = \hat{c}_{i0} + \sum_{t=1}^{n_i} \hat{c}_{it} X_{it}$, donde los coeficientes \hat{c}_{i0} y \hat{c}_{it} satisfacen las ecuaciones (2.2) y (2.3). Puesto que bajo los supuestos del modelo de Bühlmann-Straub se cumple que

$$\begin{aligned} E[\mu(\Theta_i)X_{is}] &= E[E[\mu(\Theta_i)X_{is}|\Theta_i]] = E[\mu(\Theta_i)E[X_{is}|\Theta_i]] \\ &= E[\mu(\Theta_i)^2] = Var(\mu(\Theta_i)) + E[\mu(\Theta_i)]^2 = a + m^2. \end{aligned}$$

y

$$\begin{aligned} E[X_{it}X_{is}] &= E[E[X_{it}X_{is}|\Theta_i]] = E[Cov(X_{it}X_{is}|\Theta_i) + E[X_{it}|\Theta_i]E[X_{is}|\Theta_i]] \\ &= E[\mathbf{1}_{\{t=s\}} \frac{1}{\omega_{it}} \sigma^2(\Theta_i) + \mu(\Theta_i)^2] = \mathbf{1}_{\{t=s\}} \frac{1}{\omega_{it}} s^2 + a + m^2, \end{aligned}$$

obtenemos de (2.3) que

$$\begin{aligned} a + m^2 &= m\hat{c}_{i0} + \hat{c}_{is} \left(\frac{s^2}{\omega_{is}} + a + m^2 \right) + \sum_{t \neq s} \hat{c}_{it} (a + m^2) \\ &= m\hat{c}_{i0} + \hat{c}_{is} \frac{s^2}{\omega_{is}} + \left(\sum_{t=1}^{n_i} \hat{c}_{it} \right) (a + m^2). \end{aligned}$$

Más aún, como por (2.2) $\sum_{t=1}^{n_i} \hat{c}_{it} = 1 - \frac{\hat{c}_{i0}}{m}$, se obtiene combinando resultados que

$$\hat{c}_{is} = \frac{a\omega_{is}\hat{c}_{i0}}{s^2m}.$$

Sumando todas estas cantidades y usando nuevamente (2.2), obtenemos que

$$\hat{c}_{i0} = \frac{ms^2}{a\omega_{i.} + s^2} \quad y \quad \hat{c}_{it} = \frac{a\omega_{it}}{a\omega_{i.} + s^2}.$$

En otras palabras, el estimador lineal de Bayes bajo el modelo de Bühlmann-Straub viene dado por:

$$\hat{\mu}_i^{LB} = \frac{ms^2}{a\omega_{i.} + s^2} + \sum_{t=1}^{n_i} \frac{a\omega_{it}}{a\omega_{i.} + s^2} X_{it},$$

expresión que coincide, al ordenarlo adecuadamente, con el estimador propuesto en el teorema.

De otro lado, para obtener el riesgo de Bayes notemos inicialmente que $Cov(X_{it}, X_{is}) = a + \mathbf{1}_{\{t=s\}} \frac{1}{\omega_{it}} s^2$, lo cual se obtiene directamente de los cálculos anteriores. Por tanto, según el teorema 2.1

$$\begin{aligned}\rho(\hat{\mu}_i^{LB}) &= \text{Var}(\mu(\theta_i)) - \text{Var}(\hat{\mu}_i^{LB}) \\ &= a - Z_i^2 \text{Var}(\bar{X}_{iw}).\end{aligned}$$

Operando el último término obtenemos:

$$\begin{aligned}\text{Var}(\bar{X}_{iw}) &= \frac{1}{\omega_{i.}^2} \sum_{t=1}^{n_i} \sum_{s=1}^{n_i} \left(a + 1_{\{t=s\}} \frac{1}{\omega_{it}} s^2 \right) \omega_{it} \omega_{it} \\ &= \frac{1}{\omega_{i.}^2} (a \omega_{i.}^2 + s^2 \omega_{i.}) \\ &= \frac{a \omega_{i.} + s^2}{\omega_{i.}}\end{aligned}$$

Así,

$$\begin{aligned}\rho(\hat{\mu}_i^{LB}) &= a - \frac{a^2 \omega_{i.}^2}{(a \omega_{i.} + s^2)^2} \left(\frac{a \omega_{i.} + s^2}{\omega_{i.}} \right) \\ &= a - a \left(\frac{a \omega_{i.}}{a \omega_{i.} + s^2} \right) \\ &= (1 - Z_i) a\end{aligned}$$

□

2.3.3. El modelo jerárquico de Jewell

En algunas circunstancias los portafolios puede subdividirse en sub-portafolios (o sectores) que muestren más homogeneidad entre ellos. Jewell (1975) propuso una extensión del modelo de Bühlmann-Straub considerando precisamente una partición de un portafolio de k pólizas en P sub-portafolios. Este modelo se construye jerárquicamente de tal forma que:

- (1) A cada sub-portafolio $p \in \{1, 2, \dots, P\}$ se le asocia una variable estructural Θ_p .
- (2) Cada póliza $i \in \{1, 2, \dots, k_p\}$ dentro de un sub-portafolio p está caracterizado por el par de variables estructurales (Θ_p, Θ_{pi}) . En esta póliza se pueden observar un vector de severidades $\mathbf{X}_{pi.} = (X_{pit})_{t=1,2,\dots,n_{pi}}$ asociados cada uno a los pesos ω_{pit} conocidos con antelación.

Como ejemplo podríamos considerar un portafolio constituido por un seguro de autos. En él los autos se podrían dividir en sub-portafolios constituidos por las distintas marcas de autos existentes. Más aún, esta división podría extenderse más allá de dos niveles, por citar podríamos sub-dividir estos últimos sub-portafolios en otros constituidos por los modelos dentro de cada marca de auto y así sucesivamente. Nosotros ilustraremos el modelo jerárquico considerando tan solo dos niveles. Esto lo hacemos por simplicidad, aunque su extensión a

más de dos niveles es directa en base a lo aquí estudiado.

En un modelo jerárquico de dos niveles, uno dispondrá del siguiente cuadro de severidades, distribuidos en P sub-portafolios:

Sub-portafolio	V. estructural sub-portafolio	Póliza	V. estructural póliza	Periodos				
				1	...	t	...	n_{pi}
1	Θ_1	1	Θ_{11}	X_{111}	...	X_{11t}	...	$X_{11n_{11}}$
		\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
		i	Θ_{1i}	X_{1i1}	...	X_{1it}	...	$X_{1in_{1i}}$
		\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
		k_1	Θ_{1k_1}	X_{1k_11}	...	X_{1k_1t}	...	$X_{1k_1n_{1k_1}}$
\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
p	Θ_p	1	Θ_{p1}	X_{p11}	...	X_{p1t}	...	$X_{p1n_{p1}}$
		\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
		i	Θ_{pi}	X_{pi1}	...	X_{pit}	...	$X_{pin_{pi}}$
		\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
		k_p	Θ_{pk_p}	X_{pk_p1}	...	X_{pk_pt}	...	$X_{pk_pn_{pk_p}}$
\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
P	Θ_P	1	Θ_{P1}	X_{P11}	...	X_{P1t}	...	$X_{P1n_{P1}}$
		\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
		i	Θ_{Pi}	X_{Pi1}	...	X_{Pit}	...	$X_{Pin_{Pi}}$
		\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
		k_P	Θ_{Pk_P}	X_{Pk_P1}	...	X_{Pk_Pt}	...	$X_{Pk_Pn_{Pk_P}}$

Cuadro 2.2: Modelo de Jewell

conjuntamente con una tabla de pesos conocidos ω_{pit} asociados a cada severidad X_{pit} de la tabla.

El modelo de Jewell se basa en los siguientes supuestos:

- (J_1) Los sub-portafolios $(\Theta_p, \Theta_p, \mathbb{X}_p)_{p=1,2,\dots,P}$, donde $\Theta_p = (\Theta_{pi})_{i=1,2,\dots,k_p}$ y $\mathbb{X}_p = [X_{pit}]$ es la matriz de severidades del sub-portafolio p , son independientes.
- (J_2) Dentro de cada sub-portafolio p , las pólizas $(\Theta_{pi}, \mathbb{X}_{pi})_{i=1,2,\dots,k_p}$ son condicionalmente independientes dado Θ_p .
- (J_3) Dentro de cada póliza i del sub-portafolio p , las severidades en \mathbb{X}_{pi} son independientes dado (Θ_p, Θ_{pi}) .
- (J_4) Los vectores estructurales $(\Theta_p, \Theta_{pi})_{i=1,2,\dots,n_{pk_p}}$ son para $p=1,2,\dots,P$ idénticamente distribuidos.
- (J_5) Para cualquier sub-portafolio p y póliza i :

$$E[X_{pit} | \Theta_p, \Theta_{pi}] = \mu(\Theta_p, \Theta_{pi})$$

y

$$Var(X_{pit}|\Theta_p, \Theta_{pi}) = \frac{1}{\omega_{pit}} \sigma^2(\Theta_p, \Theta_{pi}).$$

Note que el modelo de Jewell plantea una única prima de riesgo individual

$$\mu(\Theta_p, \Theta_{pi})$$

para toda póliza i al interior de un sub-portafolio p , independientemente del periodo en que se observe la severidad. Además de esta prima, será también de interés estimar la prima de riesgo del sub-portafolio p , la cual definiremos por:

$$\nu(\Theta_p) = E[\mu(\Theta_p, \Theta_{pi})|\Theta_p] = E[X_{pit}|\Theta_p].$$

Cabe reiterar que en este modelo los pesos ω_{pit} se asumen previamente dados y tienen la misma interpretación que la vista en el modelo de Bühlmann -Straub. Al igual que en este último modelo será también de utilidad considerar la suma de estos pesos al interior de cada póliza i perteneciente a un sub-portafolio p , la cual denotaremos por

$$\omega_{pi} = \sum_{t=1}^{n_{pi}} \omega_{pit}.$$

De manera análoga a los modelos previos dispondremos aquí de los siguientes parámetros estructurales en el modelo:

- $m = E[\nu(\Theta_p)] = E[\mu(\Theta_p, \Theta_{p,i})] = E[X_{pit}]$, el cual representa la prima de riesgo colectivo para todo el portafolio.
- $s^2 = E[\sigma^2(\Theta_p, \Theta_{pi})]$, el cual mide el grado de variabilidad (o heterogeneidad) media al interior de cada póliza.
- $a = E[Var(\mu(\Theta_p, \Theta_{pi})|\Theta_p)]$, el cual mide el grado de variabilidad (o heterogeneidad) media al interior de cada sub-portafolio.
- $b = Var(\nu(\Theta_p))$, el cual mide la heterogeneidad entre los diferentes sub-portafolios.

En base a los supuestos establecidos, estamos ahora en capacidad de enunciar el siguiente lema de demostración directa que nos describe la estructura de dependencia entre los distintos elementos del modelo.

Lema 2.1 Para cualesquiera $p, q = 1, 2, \dots, P$; $i = 1, 2, \dots, k_p$; $j = 1, 2, \dots, k_q$; $t = 1, 2, \dots, n_{pi}$ y $s = 1, 2, \dots, n_{qj}$

- a) $Cov(\mu(\Theta_p, \Theta_{pi}), X_{qjt}) = \mathbf{1}_{\{p=q\}}(\mathbf{1}_{\{i=j\}}a + b).$
- b) $Cov(\nu(\Theta_p), X_{qjt}) = \mathbf{1}_{\{p=q\}}b.$
- c) $Cov(X_{pit}, X_{pjs}) = \mathbf{1}_{\{i=j\}}(a + \frac{1}{\omega_{pit}}\mathbf{1}_{\{s=t\}}s^2) + b.$

Teorema 2.4 *Bajo los supuestos (J1) a (J5)*

a) *El estimador lineal de Bayes para la prima de una póliza i , dentro de un sub-portafolio p viene dado por la fórmula de credibilidad:*

$$\hat{\mu}(\Theta_p, \Theta_{pi}) = (1 - Z_{pi})m + Z_{pi}\bar{X}_{pi\omega},$$

donde

$$Z_{pi} = \frac{a\omega_{pi.}}{a\omega_{pi.} + s^2} \quad y \quad \bar{X}_{pi\omega} = \frac{1}{\omega_{pi.}} \sum_{t=1}^{n_{pi}} \omega_{pit} X_{pit}.$$

b) *El estimador lineal de Bayes para la prima de un sub-portafolio p viene dado por la fórmula de credibilidad:*

$$\hat{\nu}(\Theta_p) = (1 - Z_p)m + Z_p\bar{X}_{pz\omega},$$

donde

$$Z_p = \frac{bZ_{p.}}{bZ_{p.} + a}, \quad \bar{X}_{pz\omega} = \frac{1}{Z_{p.}} \sum_{i=1}^{k_p} Z_{pi} \bar{X}_{pi\omega} \quad y \quad Z_{p.} = \sum_{i=1}^{k_p} Z_{pi}.$$

Demostración: Puesto que el modelo de Jewell se reduce al de Bühlmann-Straub dentro de cada sub-portafolio, la parte a) es consecuencia inmediata del teorema 2.3. De otro lado, de manera análoga a lo deducido en (2.2) y (2.3) con el teorema de proyección, uno obtiene que los coeficientes óptimos para el estimador lineal de Bayes de la prima de cualquier sub-portafolio p :

$$\hat{\nu}(\Theta_p) = \hat{c}_{p0} + \sum_{i=1}^{k_p} \sum_{t=1}^{n_{pi}} \hat{c}_{pit} X_{pit},$$

deben de satisfacer las siguientes ecuaciones:

$$\hat{c}_{p0} = m \left(1 - \sum_{i=1}^{k_p} \sum_{t=1}^{n_{pi}} \hat{c}_{pit} \right)$$

y

$$E[\nu(\Theta_p) X_{pjs}] = m \hat{c}_{p0} + \sum_{i=1}^{k_p} \sum_{t=1}^{n_{pi}} \hat{c}_{pit} E[X_{pit} X_{pjs}] \dots (*),$$

para cualquier $j = 1, 2, \dots, k_p$ y $s = 1, 2, \dots, n_{pj}$. Encontremos entonces los momentos requeridos en (*). En primer lugar tenemos que

$$\begin{aligned} E[\nu(\Theta_p) X_{pjs}] &= E[E[\nu(\Theta_p) X_{pjs} | \Theta_p]] = E[\nu(\Theta_p) E[X_{pjs} | \Theta_p]] \\ &= E[\nu(\Theta_p)^2] = \text{Var}(\nu(\Theta_p)) + E[\nu(\Theta_p)]^2 = b + m^2. \end{aligned}$$

Notemos también que

$$E[\mu(\Theta_p, \Theta_{pi})^2] = E[E[\mu(\Theta_p, \Theta_{pi})^2 | \Theta_p]] = E[\text{Var}(\mu(\Theta_p, \Theta_{pi}) | \Theta_p) + E[\mu(\Theta_p, \Theta_{pi}) | \Theta_p]^2]$$

$$= a + E[\nu(\Theta_p)^2] = a + Var(\nu(\Theta_p)) + E[\nu(\Theta_p)]^2 = a + b + m^2.$$

Por otro lado, a fin de obtener $E[X_{pit}X_{pjs}]$ consideremos los casos $i = j$ e $i \neq j$. En el primero

$$\begin{aligned} E[X_{pit}X_{pjs}] &= E[X_{pit}X_{pis}] \\ &= E[E[X_{pit}X_{pis}|\Theta_p, \Theta_{pi}]] \\ &= \mathbf{1}_{\{s=t\}} E[E[X_{pit}^2|\Theta_p, \Theta_{pi}]] + \mathbf{1}_{\{s \neq t\}} E[E[X_{pit}|\Theta_p, \Theta_{pi}]E[X_{pis}|\Theta_p, \Theta_{pi}]] \\ &= \mathbf{1}_{\{s=t\}} E[Var(X_{pit}|\Theta_p, \Theta_{pi}) + E[X_{pit}|\Theta_p, \Theta_{pi}]^2] + \mathbf{1}_{\{s \neq t\}} E[\mu(\Theta_p, \Theta_{pi})^2] \\ &= \mathbf{1}_{\{s=t\}} E\left[\frac{1}{\omega_{pit}} \sigma^2(\Theta_p, \Theta_{pi}) + \mu(\Theta_p, \Theta_{pi})^2\right] + \mathbf{1}_{\{s \neq t\}} E[\mu(\Theta_p, \Theta_{pi})^2] \\ &= \mathbf{1}_{\{s=t\}} \left(\frac{s^2}{\omega_{pit}} + a + b + m^2\right) + \mathbf{1}_{\{s \neq t\}} (a + b + m^2) \\ &= a + b + m^2 + \mathbf{1}_{\{s=t\}} \frac{s^2}{\omega_{pit}}. \end{aligned}$$

En el segundo caso

$$E[X_{pit}X_{pjs}] = E[E[X_{pit}X_{pis}|\Theta_p]] = E[E[X_{pit}|\Theta_p]E[X_{pis}|\Theta_p]] = E[\nu(\Theta_p)^2] = b + m^2.$$

Consecuentemente, reemplazando en (*)

$$\begin{aligned} b + m^2 &= m \hat{c}_{p0} + \sum_{\substack{i=1 \\ i \neq j}}^{k_p} \sum_{t=1}^{n_{pi}} \hat{c}_{pit} E[X_{pit}X_{pjs}] + \sum_{\substack{t=1 \\ t \neq s}}^{n_{pi}} \hat{c}_{pjt} E[X_{pjt}X_{pjs}] + \hat{c}_{pjs} E[X_{pjs}^2] \\ &= m \hat{c}_{p0} + (b + m^2) \sum_{\substack{i=1 \\ i \neq j}}^{k_p} \sum_{t=1}^{n_{pi}} \hat{c}_{pit} + (a + b + m^2) \sum_{\substack{t=1 \\ t \neq s}}^{n_{pi}} \hat{c}_{pjt} + \hat{c}_{pjs} \left(\frac{s^2}{\omega_{pjs}} + a + b + m^2\right) \\ &= m \hat{c}_{p0} + (b + m^2) \sum_{\substack{i=1 \\ i \neq j}}^{k_p} \sum_{t=1}^{n_{pi}} \hat{c}_{pit} + (a + b + m^2) \sum_{t=1}^{n_{pj}} \hat{c}_{pjt} + \hat{c}_{pjs} \frac{s^2}{\omega_{pjs}}. \end{aligned}$$

Teniendo en cuenta que $\sum_{i=1}^{k_p} \sum_{t=1}^{n_{pi}} \hat{c}_{pit} = \sum_{\substack{i=1 \\ i \neq j}}^{k_p} \sum_{t=1}^{n_{pi}} \hat{c}_{pit} + \sum_{t=1}^{n_{pj}} \hat{c}_{pjt} = 1 - \frac{\hat{c}_{p0}}{m}$ se sigue que

$$\frac{b\hat{c}_{p0}}{m} = a \sum_{t=1}^{n_{pj}} \hat{c}_{pjt} + \hat{c}_{pjs} \frac{s^2}{\omega_{pjs}}$$

de donde

$$\hat{c}_{pjs} = \frac{1}{s^2} \left(\frac{b\hat{c}_{p0}}{m} - a \sum_{t=1}^{n_{pj}} \hat{c}_{pjt} \right) \omega_{pjs} \dots (**)$$

Sumando aquí sobre todos los periodos

$$\sum_{s=1}^{n_{pj}} \hat{c}_{pjt} = \frac{1}{s^2} \left(\frac{b\hat{c}_{p0}}{m} - a \sum_{t=1}^{n_{pj}} \hat{c}_{pjt} \right) \omega_{pj.},$$

lo cual nos da

$$\sum_{t=1}^{n_{pj}} \hat{c}_{pjt} = \frac{b\hat{c}_{p0}\omega_{pj.}}{m(a\omega_{pj.} + s^2)}.$$

En consecuencia

$$\begin{aligned} 1 - \frac{\hat{c}_{p0}}{m} &= \sum_{i=1}^{k_p} \sum_{t=1}^{n_{pi}} \hat{c}_{pit} = \frac{b\hat{c}_{p0}\omega_{pj.}}{am} \sum_{j=1}^{k_p} \frac{a\omega_{pj.}}{a\omega_{pj.} + s^2} \\ &= \frac{b\hat{c}_{p0}\omega_{pj.}}{am} \sum_{j=1}^{k_p} Z_{pj} = \frac{b\hat{c}_{p0}\omega_{pj.}}{am} Z_p. \end{aligned}$$

Despejando

$$\hat{c}_{p0} = \frac{am}{bZ_p. + a} = (1 - \frac{bZ_p.}{bZ_p. + a})m = (1 - Z_p)m.$$

Esto nos conduce en (**) a:

$$\begin{aligned} \hat{c}_{pjs} &= \frac{1}{s^2} \left(\frac{b\hat{c}_{p0}}{m} - \frac{b\hat{c}_{p0}}{m} Z_{pj} \right) \omega_{pjs} \\ &= \frac{1}{s^2} b \left(\frac{a}{bZ_p. + a} \right) \left(\frac{s^2}{a\omega_{pj.} + s^2} \right) \frac{Z_p. \omega_{pj.}}{Z_p. \omega_{pj.}} \omega_{pjs} = \frac{Z_p Z_{pj} \omega_{pjs}}{Z_p. \omega_{pj.}}. \end{aligned}$$

Reemplazando estos pesos en $\hat{\nu}(\Theta_p) = \hat{c}_{p0} + \sum_{i=1}^{k_p} \sum_{t=1}^{n_{pi}} \hat{c}_{pit} X_{pit}$ obtendremos entonces el estimador de la parte b) del teorema. \square

2.3.4. El modelo de regresión de Hachemeister

El modelo de regresión propuesto por [Hachemeister \(1975\)](#) es una extensión del modelo de Bühlmann-Straub, donde a diferencia de este la prima de riesgo de una póliza i al interior de un portafolio constituido por k pólizas es dependiente de una variable estructural Θ_i y de un conjunto de variables explicativas, lo que la hace no homogénea; es decir, dependiente también del periodo de observación a través de estas variables.

A diferencia de los modelos anteriores, el modelo de Hachemeister es multidimensional en el sentido que la prima de riesgo para un portafolio i , $\mu(\Theta_i) = E[\mathbf{X}_i | \Theta_i]$, no es una variable aleatoria sino un vector aleatorio, donde

$$\mathbf{X}_i. = [X_{i1} \ X_{i2} \ \dots \ X_{in_i}]^\top,$$

es el vector de severidades de la i -ésima póliza, el cual estamos escribiendo por conveniencia como un vector columna.

De manera general podemos definir, siguiendo a [Bühlmann y Gisler \(n.d.\)](#), el modelo de heterogeneidad multidimensional a través de los supuestos siguientes

(HM1) Las pólizas $(\Theta_i, \mathbf{X}_i.)$ se asumen independientes y las variables estructurales Θ_i son i.i.d., para $i = 1, 2, \dots, k$.

(HM2) Para cualquier $i = 1, 2, \dots, k$:

$$\text{Cov}(\mathbf{X}_i | \Theta_i) = \Sigma(\Theta_i)$$

es una matriz simétrica positiva semi-definida.

Siendo nuestro interés estimar $\boldsymbol{\mu}(\Theta_i)$ basaremos esta estimación en los siguientes parámetros estructurales:

- $\boldsymbol{\mu} = E[\boldsymbol{\mu}(\Theta_i)]$.
- $\mathbb{A} = \text{Cov}(\boldsymbol{\mu}(\Theta_i))$.
- $\mathbb{S}_i = E[\Sigma(\Theta_i)]$.

El teorema siguiente nos brinda la forma del estimador de credibilidad lineal de Bayes para el modelo de heterogeneidad multidimensional.

Teorema 2.5 *Bajo los supuestos (HM1) y (HM2) el estimador lineal de Bayes para la prima de riesgo de la póliza i viene dado por la fórmula de credibilidad:*

$$\hat{\mu}(\Theta_i) = (\mathbb{I} - \mathbb{Z}_i)\boldsymbol{\mu} + \mathbb{Z}_i\mathbf{X}_i,$$

donde

$$\mathbb{Z}_i = \mathbb{A}(\mathbb{A} + \mathbb{S}_i)^{-1}.$$

Demostración: Por el supuesto de independencia entre pólizas bastará trabajar al interior de cualquiera de ellas; así consideraremos en adelante una póliza i arbitraria y suprimiremos por simplicidad el subíndice i cuando sea conveniente.

Consideremos el sub-espacio cerrado de L^2 de todas las variables aleatorias que sean funciones afín-lineales de las componentes del vector $\mathbf{X} \equiv \mathbf{X}_i$. El teorema de proyección garantiza la existencia de v.a's únicas en este espacio que minimizan la media cuadrática entre las componentes $\mu_t(\Theta_i)$ de $\boldsymbol{\mu}(\Theta_i)$ y estas proyecciones. Denotemos pues, para cualquier $t = 1, 2, \dots, n_i$, por

$$\hat{\mu}_t(\Theta_i) = \hat{c}_{t0} + \sum_{s=1}^{n_i} \hat{c}_{ts} X_s$$

a tales proyecciones. Puesto que por el mismo teorema de proyección $\mu_t(\Theta_i) - \hat{\mu}_t(\Theta_i)$ es ortogonal a toda combinación afín-lineal de las componentes de \mathbf{X} , lo es en particular para cualquier constante y también para las variables $X_s - \mu_s$. Así:

$$\mu_t = \hat{c}_{t0} + \sum_{s=1}^{n_i} \hat{c}_{ts} \mu_s$$

y

$$\text{Cov}(\mu_t(\Theta_i), X_s) = \sum_{r=1}^{n_i} \hat{c}_{tr} \text{Cov}(X_r, X_s).$$

Expresando estas ecuaciones matricialmente obtenemos:

$$\boldsymbol{\mu} = \hat{\mathbf{c}}_0 + \hat{\mathbf{C}}\boldsymbol{\mu}$$

y

$$\text{Cov}(\boldsymbol{\mu}(\Theta_i), \mathbf{X}) = \hat{\mathbf{C}} \text{Cov}(\mathbf{X}),$$

donde $\hat{\mathbf{C}}_0 = [c_{10} \ c_{20} \ \dots \ c_{n_i 0}]^T$ y $\hat{\mathbf{C}} = [\hat{c}_{ts}]$.

Más aún,

$$\begin{aligned} \text{Cov}(\boldsymbol{\mu}(\Theta_i), \mathbf{X}) &= E[\text{Cov}(\boldsymbol{\mu}(\Theta_i), \mathbf{X}|\Theta_i)] + \text{Cov}(\boldsymbol{\mu}(\Theta_i), E[\mathbf{X}|\Theta_i]) \\ &= \text{Cov}(\boldsymbol{\mu}(\Theta_i), \boldsymbol{\mu}(\Theta_i)) = \text{Cov}(\boldsymbol{\mu}(\Theta_i)) = \mathbb{A} \end{aligned}$$

y

$$\text{Cov}(\mathbf{X}) = E[\text{Cov}(\mathbf{X}, \mathbf{X}|\Theta_i)] + \text{Cov}(E[\mathbf{X}|\Theta_i], E[\mathbf{X}|\Theta_i]) = E[\Sigma(\Theta_i)] + \text{Cov}(\boldsymbol{\mu}(\Theta_i)) = (\mathbb{S}_i + \mathbb{A}).$$

Por tanto, de reemplazarse estas expresiones en la ecuación anterior se obtiene que

$$\mathbb{A} = \hat{\mathbf{C}}(\mathbb{A} + \mathbb{S}_i),$$

ó $\hat{\mathbf{C}} = \mathbb{A}(\mathbb{A} + \mathbb{S}_i)^{-1} = \mathbb{Z}_i$ y seguidamente que $\hat{\mathbf{c}}_0 = (\mathbb{I} - \mathbb{Z}_i)\boldsymbol{\mu}$. Esto nos lleva directamente al estimador lineal de Bayes. \square

Introduzcamos ahora el modelo de Hachemeister como un modelo particular de heterogeneidad multidimensional en el que el vector de primas para una póliza i puede ser explicado en cada periodo por un conjunto de q variables no aleatorias a través del sistema de ecuaciones de regresión:

$$\begin{aligned} \mu_1(\Theta_i) &= E[X_{i1}|\Theta_i] = \beta_0(\Theta_i) + y_{11}\beta_1(\Theta_i) + \dots + y_{1h}\beta_h(\Theta_i) + \dots + y_{1q}\beta_q(\Theta_i) \\ \mu_2(\Theta_i) &= E[X_{i2}|\Theta_i] = \beta_0(\Theta_i) + y_{21}\beta_1(\Theta_i) + \dots + y_{2h}\beta_h(\Theta_i) + \dots + y_{2q}\beta_q(\Theta_i) \\ &\vdots \\ \mu_t(\Theta_i) &= E[X_{it}|\Theta_i] = \beta_0(\Theta_i) + y_{t1}\beta_1(\Theta_i) + \dots + y_{th}\beta_h(\Theta_i) + \dots + y_{tq}\beta_q(\Theta_i) \\ &\vdots \\ \mu_{n_i}(\Theta_i) &= E[X_{in_i}|\Theta_i] = \beta_0(\Theta_i) + y_{n_i 1}\beta_1(\Theta_i) + \dots + y_{n_i h}\beta_h(\Theta_i) + \dots + y_{n_i q}\beta_q(\Theta_i) \end{aligned}$$

ó matricialmente como

$$\boldsymbol{\mu}(\Theta_i) = E[\mathbf{X}_i|\Theta_i] = \mathbb{Y}\boldsymbol{\beta}(\Theta_i), \quad (2.4)$$

donde $\mathbb{Y} = [y_{th}]$ es una matriz $n_i \times (q+1)$ de variables explicativas conocidas con un vector columna de unos al inicio y $\boldsymbol{\beta}(\Theta_i)$ es un vector aleatorio columna $(q+1) \times 1$ de coeficientes estructurales de regresión para la póliza i .

Siendo este un modelo de regresión es menester dotarlo de una estructura de covarianza. En general se plantea que este puede ser un modelo heterocedástico. Más explícitamente, el modelo de Hachemeister asume los supuestos siguientes

(H1) Las pólizas (Θ_i, \mathbf{X}_i) se asumen independientes y las variables estructurales Θ_i son i.i.d. para $i = 1, 2, \dots, k$.

(H2) Se cumple (2.4), donde \mathbb{Y} es una matriz de rango completo.

(H3) La matriz de varianza-covarianza en (2.4) viene dada por

$$\text{Cov}(\mathbf{X}_i | \Theta_i) = \Sigma(\Theta_i),$$

donde esta última matriz podría depender de ciertos pesos asociados a cada severidad y de otro u otros parámetros.

Una forma muy común de especificar la matriz de varianza-covarianza en (2.4) es por ejemplo tomando:

$$\Sigma(\Theta_i) = \sigma^2(\Theta_i) \mathbb{V}_i, \quad (2.5)$$

siendo $\sigma^2(\Theta_i)$ un parámetro estructural unidimensional y \mathbb{V}_i una matriz $n_i \times n_i$ simétrica y positiva semi-definida de pesos inversos conocidos, como en el modelo de Bühlmann-Straub.

El modelo de Hachemeister incorpora los siguientes parámetros estructurales

- $\mathbf{b} = E[\beta(\Theta_i)]$.
- $\mathbb{A} = \text{Cov}(\beta(\Theta_i))$.
- $\mathbb{S}_i = E[\text{Cov}(\mathbf{X}_i | \Theta_i)]$.

Dados los supuestos de este modelo uno podría pensar en que si Θ_i fuese dado, entonces el vector de coeficientes de regresión podría estimarse mediante el estimador de mínimos cuadrados generalizados que ilustraremos en el próximo capítulo; sin embargo este no es en sí un estimador pues depende de la matriz de varianza-covarianza $\Sigma(\Theta_i)$, el cual en general podría involucrar otros parámetros. En tal sentido, una manera intuitiva de solucionar esto sería reemplazándola por su parámetro estructural \mathbb{S}_i . La siguiente proposición confirma precisamente esta sospecha.

Proposición 2.1 *El mejor estimador lineal y condicionalmente insesgado de $\beta(\Theta_i)$, basado en \mathbf{X}_i , viene dado por:*

$$\mathbf{B}_i = (\mathbb{Y}^\top \mathbb{S}_i^{-1} \mathbb{Y})^{-1} \mathbb{Y}^\top \mathbb{S}_i^{-1} \mathbf{X}_i.$$

Demostración: Dado que $\beta(\Theta_i)$ es un vector aleatorio que depende de la variable estructural Θ_i , es preciso aclarar que definimos aquí al mejor estimador lineal y condicionalmente insesgado de $\beta(\Theta_i)$, basado en \mathbf{X}_i , como un estimador de la forma $\mathbb{C}\mathbf{X}_i$ que satisfaga $E[\mathbb{C}\mathbf{X}_i | \Theta_i] = \beta(\Theta_i)$ y que haga que

$$\text{Cov}(\mathbb{C}\mathbf{X}_i)$$

sea mínima en el sentido que la diferencia de esta matriz de varianza-covarianza con otra que satisfaga las condiciones anteriores sea positiva semi-definida.

Observemos primero que la matriz de varianza-covarianza anterior, puede escribirse por condicionamiento como:

$$\begin{aligned} Cov(\mathbb{C}\mathbf{X}_i) &= E[Cov(\mathbb{C}\mathbf{X}_i|\Theta_i)] + Cov(E[\mathbb{C}\mathbf{X}_i|\Theta_i]) \\ &= \mathbb{C}E[Cov(\mathbf{X}_i|\Theta_i)]\mathbb{C}^\top + Cov(\beta(\Theta_i)) = \mathbb{C}E[\Sigma(\Theta_i)]\mathbb{C}^\top + \mathbb{A} = \mathbb{C}\mathbb{S}_i\mathbb{C}^\top + \mathbb{A}. \end{aligned}$$

En tal sentido, la minimización de $Cov(\mathbb{C}\mathbf{X}_i)$ equivale a la minimización de $\mathbb{C}\mathbb{S}_i\mathbb{C}^\top$ sobre todas las matrices \mathbb{C} que satisfagan $E[\mathbb{C}\mathbf{X}_i|\Theta_i] = \beta(\Theta_i)$. Esto es precisamente (3.3) con $\Sigma_i = \mathbb{S}_i$, de lo cual se sigue directamente la proposición. \square

Teorema 2.6 *Bajo los supuestos (H1) a (H3) el estimador lineal de Bayes para el vector de coeficientes de regresión estructurales en la póliza i viene dado por la fórmula de credibilidad:*

$$\hat{\beta}(\Theta_i) = (\mathbb{I} - \mathbb{Z}_i)\mathbf{b} + \mathbb{Z}_i\mathbf{B}_i$$

donde

$$\mathbb{Z}_i = \mathbb{A}(\mathbb{A} + \mathbb{U}_i)^{-1} \quad y \quad \mathbb{U}_i = (\mathbb{Y}^\top \mathbb{S}_i^{-1} \mathbb{Y})^{-1}.$$

Demostración: Esto es una aplicación de los dos últimos resultados utilizando el modelo de heterogeneidad multidimensional no aplicado a la prima de riesgo de un portafolio i y su estimación mediante el vector de severidades \mathbf{X}_i , sino aplicados a su vector de coeficientes de regresión estructurales y su estimación mediante los $\mathbf{B}_i = \mathbb{H}_i\mathbf{X}_i$, donde $\mathbb{H}_i = (\mathbb{Y}^\top \mathbb{S}_i^{-1} \mathbb{Y})^{-1} \mathbb{Y}^\top \mathbb{S}_i^{-1}$.

En efecto, si nos interesa obtener el estimador lineal de Bayes de $\beta(\Theta_i) = E[\mathbf{B}_i|\Theta_i]$, expresión última valedera por la proposición anterior, entonces vemos fácilmente que por los supuesto (H1) a (H3) se satisfacen también (HM1) y (HM2) de tomarse \mathbf{B}_i en lugar de \mathbf{X}_i . En este contexto

$$Cov(\mathbf{B}_i|\Theta_i) = \mathbb{H}_i \Sigma(\Theta_i) \mathbb{H}_i^\top \equiv \tilde{\Sigma}(\Theta_i)$$

sigue siendo una matriz simétrica positiva semi-definida, y se tienen los parámetros estructurales: $\mathbf{b} = E[\beta(\Theta_i)]$, $\mathbb{A} = Cov(\beta(\Theta_i))$ y

$$\mathbb{U}_i = E[\tilde{\Sigma}(\Theta_i)] = \mathbb{H}_i E[\Sigma(\Theta_i)] \mathbb{H}_i^\top = \mathbb{H}_i \mathbb{S}_i \mathbb{H}_i^\top = (\mathbb{Y}^\top \mathbb{S}_i^{-1} \mathbb{Y})^{-1}.$$

Así, este teorema se sigue de una aplicación directa del teorema 2.5. \square

Una vez estimado el vector de coeficientes de regresión estructurales de una póliza i , queda simplemente estimar la prima de Bayes de ella para un vector columna \mathbf{Y} de orden $q + 1$ de variables explicativas, mediante

$$\hat{\mu}(\Theta_i) = \mathbf{Y} \hat{\beta}(\Theta_i). \quad (2.6)$$

2.4. Estimación clásica de los parámetros estructurales

2.4.1. Estimadores para el modelo Jerárquico

Puesto que el modelo jerárquico de Jewell involucra como casos particulares a los de Bühlmann y Bühlmann-Straub, consideraremos directamente la estimación de este modelo y veremos los últimos como casos especiales.

Para estimar los parámetros estructurales m, b, a y s^2 (Goovaerts y Hoogstad, 1987) sugirieron el siguiente conjunto de pseudoestimadores

$$\begin{aligned}\hat{s}^2 &= \left(\sum_{p=1}^P \sum_{i=1}^{k_p} \sum_{t=1}^{n_{pi}} \omega_{pit} (X_{pit} - \bar{X}_{pi\omega})^2 \right) \left(\sum_{p=1}^P \sum_{i=1}^{k_p} (n_{pi} - 1) \right)^{-1} \\ \hat{a} &= \left(\sum_{p=1}^P \sum_{i=1}^{k_p} Z_{pi} (\bar{X}_{pi\omega} - \bar{X}_{pz\omega})^2 \right) \left(\sum_{p=1}^P (k_p - 1) \right)^{-1} \\ \hat{b} &= \left(\sum_{p=1}^P Z_p (\bar{X}_{pz\omega} - \bar{X}_{zz\omega})^2 \right) (P - 1)^{-1},\end{aligned}$$

donde $\bar{X}_{zz\omega} = \sum_{p=1}^P \frac{Z_p}{Z_{..}} \bar{X}_{pz\omega}$ y $Z_{..} = \sum_{p=1}^P Z_p$. El estimador de la prima riesgo colectivo m viene además dado por $\hat{m} = \bar{X}_{zz\omega}$.

Para poder calcular \hat{s}^2 no existe ningún problema, ya que este depende de las severidades a observar X_{pit} y de la media ponderada para cada póliza $\bar{X}_{pi\omega}$; sin embargo, para hallar \hat{a} y \hat{b} es necesario calcular primero los factores de credibilidad Z_{pi} y Z_p que dependen también de a y b . Estimadores de este tipo solo se resuelven por métodos iterativos. Esto es, escogemos un valor inicial para \hat{a} , con éste calculamos unos valores para Z_{pi} y $\bar{X}_{pz\omega}$, con estos valores encontraremos nuevos valores para \hat{a} y así sucesivamente hasta que el valor de \hat{a} se estabilice.

Luego se procede de la misma manera para \hat{b} , dándole un valor inicial y utilizando el valor de \hat{a} calculado anteriormente para hallar Z_p y $\bar{X}_{zz\omega}$, con ellos encontramos nuevos valores de \hat{b} y así sucesivamente hasta que el valor de \hat{b} se estabilice.

La utilización de un proceso iterativo proporciona una solución, que sin embargo en algunas ocasiones converge a cero. Así, por ejemplo, si $\hat{a} = 0$, indicaría que no hay heterogeneidad dentro de las subportafolios. Pons (1991) sugiere pensar por ejemplo en otra subdivisión para los subportafolios. Una mejor solución consiste en utilizar otros nuevos estimadores como los propuestos por Bühlmann y Gisler (n.d.) y Ohlsson (2006). El estimador de s^2 para ambos autores es el mismo anterior; pero los estimadores de a y b se basan en

$$\begin{aligned}A_p &= \sum_{i=1}^{k_p} \omega_{pi} (\bar{X}_{pi\omega} - \bar{X}_{p\omega\omega})^2 - (k_p - 1) \hat{s}^2 \\ B &= \sum_{p=1}^P Z_p (\bar{X}_{pz\omega} - \bar{X}_{zz\omega})^2 - (P - 1) a\end{aligned}$$

y

$$\bar{X}_{p\omega\omega} = \sum_{i=1}^{k_p} \frac{\omega_{pi.}}{\omega_{p..}} \bar{X}_{pi\omega}, \text{ con } \omega_{p..} = \sum_{i=1}^{k_p} \omega_{pi.}.$$

Por un lado [Bühlmann y Gisler \(n.d.\)](#) proponen los estimadores

$$\hat{a} = \frac{1}{P} \sum_{p=1}^P \max\left\{\frac{A_p}{c_p}, 0\right\} \quad \text{y} \quad \hat{b} = \max\left\{\frac{B}{d}, 0\right\},$$

donde B se calcula con \hat{a} ; mientras que [Ohlsson \(2006\)](#) propone los estimadores

$$\hat{a} = \frac{\sum_{p=1}^P A_p}{\sum_{p=1}^P c_p} \quad \text{y} \quad \hat{b} = \frac{B}{d},$$

siendo $c_p = \omega_{p..} - \sum_{i=1}^{k_p} \frac{\omega_{pi.}^2}{\omega_{p..}}$ y $d = Z_{..} - \sum_{p=1}^P \frac{Z_{p.}^2}{Z_{..}}$.

Un estudio comparativo entre estos estimadores ha sido realizado por [Dubey \(1981\)](#) y [Ohlsson \(2006\)](#). Ninguno de ellos se prueba que es universalmente mejor que los otros, por lo que en nuestro estudio adoptaremos el estimador de Bühlmann y Gisler.

Como ya se mencionó el modelo jerárquico de Jewell involucra como casos particulares a los de Bühlmann y Bühlmann-Straub. Estos son encontrados fácilmente si en las formulas anteriores se considera un único portafolio, sin divisiones; es decir, con $P = 1$. Más aún, como el modelo de Bühlmann es un caso particular del de Bühlmann-Straub con pesos unitarios consideraremos sin pérdida de generalidad sólo este último caso. En este modelo el número de pólizas es igual a k y el número de observaciones por póliza viene dado por n_i . En tal sentido la fórmula para \hat{s}^2 se reduce a:

$$\hat{s}^2 = \left(\sum_{i=1}^k \sum_{t=1}^{n_i} \omega_{it} (X_{it} - \bar{X}_{i\omega})^2 \right) \left(\sum_{i=1}^k (n_i - 1) \right)^{-1}$$

y similar las ecuaciones para A_p y c_p se reducen a:

$$c = \omega_{..} - \sum_{i=1}^k \frac{\omega_{i.}^2}{\omega_{..}}$$

y

$$A = \sum_{i=1}^k \omega_{i.} (\bar{X}_{i\omega} - \bar{X}_{\omega\omega})^2 - (k - 1) \hat{s}^2,$$

siendo $\bar{X}_{\omega\omega} = \frac{1}{\omega_{..}} \sum_{i=1}^k \omega_{i.} \bar{X}_{i\omega}$ y $\omega_{..} = \sum_{i=1}^k \omega_{i.}$. Así $\hat{a} = \frac{A}{c}$ es el estimador insesgado de Ohlsson y $\max\{\hat{a}, 0\}$ el de Bühlmann y Gisler.

2.4.2. Estimadores del modelo de regresión de Hachemeister

Consideremos, sin pérdida de generalidad, la especificación (2.5) para la estructura de covarianza en este modelo. Bajo ella el parámetro estructural $\mathbb{S}_i = E[\Sigma(\Theta_i)]$ se reduce a $\mathbb{S}_i = s^2 \mathbb{V}_i$, donde $s^2 = E[\sigma^2(\Theta_i)]$ es un parámetro estructural unidimensional de volatilidad

en el modelo y \mathbb{V}_i una matriz conocida. El estimador insesgado natural entonces para este parámetro viene dado por el promedio de las estimaciones en el modelo heterocedástico (2.4) de los errores estándares de estimación en cada póliza; vale decir por:

$$\hat{s}^2 = \frac{1}{k} \sum_{i=1}^k \frac{1}{n_i - q - 1} (\mathbf{X}_{i.} - \mathbb{Y} \hat{\beta}(\Theta_i))^\top \mathbb{V}_i^{-1} (\mathbf{X}_{i.} - \mathbb{Y} \hat{\beta}(\Theta_i)).$$

(Goovaerts et al., 1990) muestran por otro lado que los siguientes son estimadores insesgados para los parámetros estructurales \mathbf{b} y \mathbb{A} , respectivamente:

$$\hat{\mathbf{b}} = \left(\sum_{i=1}^k \mathbb{Z}_i \right)^{-1} \sum_{i=1}^k \mathbb{Z}_i \hat{\beta}(\Theta_i)$$

y

$$\hat{\mathbb{A}} = \frac{1}{k-1} \sum_{i=1}^k \mathbb{Z}_i (\hat{\beta}(\Theta_i) - \hat{\mathbf{b}})(\hat{\beta}(\Theta_i) - \hat{\mathbf{b}})^\top.$$

Al igual que en el modelo de Jewell las dos últimas expresiones nos proveen formalmente de pseudo-estimadores de \mathbf{b} y \mathbb{A} , pues ellos dependen de la matriz de credibilidad, la cual depende a su vez de los parámetros estructurales. Estos entonces deben de encontrarse con métodos iterativos, empezando por una estimación inicial de la matriz \mathbb{A} .

2.5. Aplicación con datos de Hachemeister

2.5.1. Presentación de los datos

En esta sección evaluaremos, a manera de ilustración, las primas de credibilidad bajo los estimadores clásicos de los parámetros estructurales vistos en la sección anterior. Una vez estimados estos parámetros estructurales serán reemplazados en las fórmulas de las primas o estimadores lineales de Bayes obteniéndose los denominados estimadores de primas de credibilidad empíricos. Utilizaremos con este fin los datos propuestos por Hachemeister (1975), el cual considera cinco pólizas (correspondientes a 5 estados de los Estados Unidos) y doce trimestres de experiencia de reclamaciones. Esta experiencia consiste en montos promedio de reclamos de pasajeros asegurados para un seguro de responsabilidad civil que sufrieron accidentes entre Julio de 1970 hasta Junio de 1973; es decir, tenemos aquí $k = 5$ pólizas y $n_i = n = 12$ periodos de observación. Los datos de estas severidades, X_{it} , se muestran a continuación.

Pólizas	t=1	t=2	t=3	t=4	t=5	t=6	t=7	t=8	t=9	t=10	t=11	t=12
1	1,738	1,642	1,794	2,051	2,079	2,234	2,032	2,035	2,115	2,262	2,267	2,517
2	1,364	1,408	1,597	1,444	1,342	1,675	1,470	1,448	1,464	1,831	1,612	1,471
3	1,759	1,685	1,479	1,763	1,674	2,103	1,502	1,622	1,828	2,155	2,233	2,059
4	1,223	1,146	1,010	1,257	1,426	1,532	1,953	1,123	1,343	1,243	1,762	1,306
5	1,456	1,499	1,609	1,741	1,482	1,572	1,606	1,735	1,607	1,573	1,613	1,690

Cuadro 2.3: Data de Hachemeister para el número promedio de reclamaciones de 5 pólizas en 12 periodos (Hachemeister, 1975)

Asociados a los promedios anteriores se tienen también las cantidad de reclamos, sobre los cuales se tomaron estos promedios. Estos constituirán los pesos en nuestra aplicación y se ilustran seguidamente.

Pólizas	t=1	t=2	t=3	t=4	t=5	t=6	t=7	t=8	t=9	t=10	t=11	t=12
1	7,861	9,251	8,706	8,575	7,917	8,263	9,456	8,003	7,365	7,832	7,849	9,077
2	1,622	1,742	1,523	1,515	1,622	1,602	1,964	1,515	1,527	1,748	1,654	1,861
3	1,147	1,357	1,329	1,204	998	1,077	1,277	1,218	896	1,003	1,108	1,121
4	407	396	348	341	315	328	352	331	287	384	321	342
5	2,902	3,172	3,046	3,068	2,693	2,910	3,275	2,697	2,663	3,017	3,242	3,425

Cuadro 2.4: Pesos asociados a los datos del Cuadro 2.3 (Hachemeister, 1975).

Para el cálculo de las estimaciones procederemos de manera manual utilizando los estimadores aquí vistos; aunque como lo explicaremos en el capítulo de simulación, podríamos también haber obtenido idénticos resultados usado el paquete actuar de R que tiene ya implementadas estas estimaciones.

2.5.2. Resultados utilizando el modelo de Bühlmann

Con los datos del cuadro 2.3, podemos encontrar las siguientes estimaciones clásicas de los parámetros estructurales para el modelo de Bühlmann:

$$\hat{m} = \bar{x} = \frac{1}{kn} \sum_{t=1}^n \sum_{i=1}^k x_{it} = \frac{100,261}{(5)(12)} = 1,671$$

$$\hat{s}^2 = \frac{1}{k(n-1)} \sum_{t=1}^n \sum_{i=1}^k (x_{it} - \bar{x}_i)^2$$

$$\hat{s}^2 = \frac{(1,738 - 2,064)^2 + \dots + (2,517 - 2,064)^2 + \dots + (1,690 - 1,599)^2}{5(12-1)} = 46,040.5$$

$$\hat{a} = \frac{1}{(k-1)} \sum_{i=1}^k (\bar{x}_i - \bar{x})^2 - \frac{\hat{s}^2}{n} = \frac{(2,064 - 1,671)^2 + \dots + (1,599 - 1,671)^2}{4} - \frac{46,040.5}{12} = 72,310$$

$$\hat{Z}_i = \hat{Z} = \frac{n\hat{a}}{\hat{s}^2 + n\hat{a}}$$

$$\hat{Z} = \frac{12(72,310)}{46,040.5 + 12(72,310)} = 0.9496$$

La prima de credibilidad empírica para cada póliza puede entonces ser encontrada mediante la fórmula:

$$\hat{\mu}(\Theta_i) = (1 - \hat{Z})\hat{m} + \hat{Z}\bar{x}_i.$$

Así tenemos:

$$\hat{\mu}(\Theta_1) = (1 - 0.9496)1,671 + 0.9496(2,064) = 2,044$$

$$\hat{\mu}(\Theta_2) = (1 - 0.9496)1,671 + 0.9496(1,511) = 1,519$$

$$\hat{\mu}(\Theta_3) = (1 - 0.9496)1,671 + 0.9496(1,822) = 1,814$$

$$\hat{\mu}(\Theta_4) = (1 - 0.9496)1,671 + 0.9496(1,360) = 1,376$$

$$\hat{\mu}(\Theta_5) = (1 - 0.9496)1,671 + 0.9496(1,599) = 1,602.$$

2.5.3. Resultados utilizando el modelo de Bühlmann-Straub

Aquí la póliza i esta definida por $(\Theta_i, x_{i1}, \dots, x_{i12})$, pero para cada severidad existe un peso asociado ω_{it} , cuyas sumas por póliza se muestran en el cuadro 2.5.

	suma de pesos
$\omega_{1.}$	100,155
$\omega_{2.}$	19,895
$\omega_{3.}$	13,735
$\omega_{4.}$	4,152
$\omega_{5.}$	36,110
$\omega_{..}$	174,047

Cuadro 2.5: Suma de pesos

Los promedios ponderados para cada póliza son:

$$\begin{aligned} \bar{x}_{1\omega} &= \frac{1,738*7,861+1,642*9,251+\dots+2,517*9,077}{100,155} = 2,061 \\ \bar{x}_{2\omega} &= \frac{1,364*1,622+1,408*1,742+\dots+1,471*1,861}{19,895} = 1,511 \\ \bar{x}_{3\omega} &= \frac{1,759*1,147+1,685*1,357+\dots+2,059*1,121}{13,735} = 1,806 \\ \bar{x}_{4\omega} &= \frac{1,223*407+1,146*396+\dots+1,306*342}{4,152} = 1,353 \\ \bar{x}_{5\omega} &= \frac{1,456*2,902+1,499*3,172+\dots+1,690*3,425}{36,110} = 1,600 \end{aligned}$$

Los siguientes son las estimaciones insesgadas de los parámetros estructurales:

$$\hat{s}^2 = \left(\frac{1}{k(n-1)} \right) \sum_{i=1}^k \sum_{t=1}^n \omega_{it} (x_{it} - \bar{x}_{i\omega})^2 = 139'120,025.9$$

$$\hat{a} = \omega_{..} \frac{\sum_{i=1}^k \omega_{i.} (\bar{x}_{i\omega} - \bar{x}_{\omega\omega})^2 - (k-1)\hat{s}^2}{\omega_{..}^2 - \sum_{i=1}^k \omega_{i.}^2} = 89,638.7$$

$$\hat{m} = 1,683.7$$

El factor de credibilidad estimado para cada contrato es:

$$\hat{Z}_1 = 0.9847 ; \hat{Z}_2 = 0.9276 ; \hat{Z}_3 = 0.8985 ; \hat{Z}_4 = 0.7279 ; \hat{Z}_5 = 0.9588$$

La prima de credibilidad empírica $\mu(\Theta_i)$ viene dada por:

$$\hat{\mu}(\Theta_i) = \hat{m} + \hat{Z}_i(\bar{x}_{i\omega} - \hat{m}),$$

Así tenemos:

$$\hat{\mu}(\Theta_1) = 1,683,7 + 0.9847(2,061 - 1,683,7) = 2,055$$

$$\hat{\mu}(\Theta_2) = 1,683,7 + 0.9276(1,511 - 1,683,7) = 1,524$$

$$\hat{\mu}(\Theta_3) = 1,683,7 + 0.8985(1,806 - 1,683,7) = 1,793$$

$$\hat{\mu}(\Theta_4) = 1,683,7 + 0.7279(1,353 - 1,683,7) = 1,443$$

$$\hat{\mu}(\Theta_5) = 1,683,7 + 0.9588(1,600 - 1,683,7) = 1,603$$

Al comparar estos resultados con los encontrados por el modelo de Bühlmann tenemos:

Póliza	Promedio para cada contrato \bar{x}_i	Prima empírica para el modelo de Bühlmann $\hat{\mu}(\Theta_i)$	Prima empírica para el modelo de Bühlmann-Straub $\hat{\mu}(\Theta_i)$
1	2,064	2,044	2,055
2	1,511	1,519	1,524
3	1,822	1,814	1,793
4	1,360	1,376	1,443
5	1,599	1,602	1,603

Cuadro 2.6: Estimaciones de primas obtenidas utilizando el modelo de Bühlmann y el modelo de Bühlmann-Straub

Recordamos que el modelo de Bühlmann-Straub es una extensión del modelo de Bühlmann, el cual tenía la limitación de no utilizar información muy importante, como es en este caso el número de reclamos por póliza. Como lo menciona Pons (1991) ambos modelos sólo son adecuados cuando se dispone de datos sin tendencia, pero además para poder aplicar el modelo de Bühlmann es necesario disponer de una cartera formada por pólizas o contratos que tenga más o menos igual importancia dentro de la misma.

2.5.4. Resultados utilizando el modelo de jerárquico de Jewell

Para ilustrar el modelo jerárquico de Jewell, subdividiremos las cinco pólizas de los datos de Hachemeister en $P=2$ sub-portafolios de la manera siguiente:

- Sub-portafolio 1: contratos 1 y 3
- Sub-portafolio 2: contratos 2, 4 y 5.

Los datos quedan entonces distribuidos de la siguiente manera:

Sub-portafolio 1												
Pólizas	t=1	t=2	t=3	t=4	t=5	t=6	t=7	t=8	t=9	t=10	t=11	t=12
1	1,738	1,642	1,794	2,051	2,079	2,234	2,032	2,035	2,115	2,262	2,267	2,517
3	1,759	1,685	1,479	1,763	1,674	2,103	1,502	1,622	1,828	2,155	2,233	2,059
Sub-portafolio 2												
2	1,364	1,408	1,597	1,444	1,342	1,675	1,470	1,448	1,464	1,831	1,612	1,471
4	1,223	1,146	1,010	1,257	1,426	1,532	1,953	1,123	1,343	1,243	1,762	1,306
5	1,456	1,499	1,609	1,741	1,482	1,572	1,606	1,735	1,607	1,573	1,613	1,690

Cuadro 2.7: Datos de Hachemeister (1975) clasificados en 2 sub-portafolios

Utilizando las ecuaciones vistas podemos encontrar:

Promedio ponderado para la póliza 1 y sub-portafolio 1: $\bar{x}_{11\omega} = 2,061$

Promedio ponderado para la póliza 3 y sub-portafolio 1: $\bar{x}_{13\omega} = 1,806$

Promedio ponderado para la póliza 2 y sub-portafolio 2: $\bar{x}_{22\omega} = 1,511$

Promedio ponderado para la póliza 4 y sub-portafolio 2: $\bar{x}_{24\omega} = 1,353$

Promedio ponderado para la póliza 5 y sub-portafolio 2: $\bar{x}_{25\omega} = 1,600$

De aquí encontramos que

$$\hat{s}^2 = 139'120,025.9$$

y utilizando las estimaciones de Bühlmann y Gesler para \hat{a} y \hat{b} obtenemos las primas de credibilidad empíricas:

$$\hat{\mu}(\Theta_1, \Theta_{11}) = 2,035$$

$$\hat{\mu}(\Theta_1, \Theta_{13}) = 1,997$$

$$\hat{\mu}(\Theta_2, \Theta_{22}) = 1,544$$

$$\hat{\mu}(\Theta_2, \Theta_{24}) = 1,544$$

$$\hat{\mu}(\Theta_2, \Theta_{25}) = 1,560$$

Resumiendo la información y comparando los resultados con el modelo de Bühlmann-Straub tenemos:

Póliza	Promedio para cada contrato \bar{X}_i	Prima empírica para el modelo de Bühlmann-Straub $\hat{\mu}(\Theta_i)$	Prima empírica para el modelo de Jewell $\hat{\mu}(\Theta_p; \Theta_{pi})$
1	2,064	2,055	2,035
2	1,511	1,524	1,544
3	1,822	1,793	1,997
4	1,360	1,443	1,544
5	1,599	1,603	1,560

Cuadro 2.8: Resultados obtenidos utilizando el Modelo de Bühlmann-Straub y el Modelo de Jewell

Observemos que el valor de $\hat{s}^2 = 139'120,025.9$ coincide para el modelo de Bühlmann-Straub y el modelo de Jewell. Esto se debe a que no se ha omitido información y que todos los contratos cuentan con el mismo número de datos.

2.5.5. Resultados utilizando el modelo de regresión de Hachemeister

Seguiremos aquí el enfoque de [Hachemeister \(1975\)](#) quién propuso para su data un modelo con tendencia lineal, tomando como covariable el tiempo; es decir, proponiendo un modelo de la forma

$$\mu_t(\Theta_i) = E[\mathbf{X}_{it}|\Theta_i] = \beta_0(\Theta_i) + t\beta_1(\Theta_i), \quad t = 1, 2, \dots, 12.$$

En este contexto la matriz \mathbb{Y} en (2.4) se expresa como:

$$\mathbb{Y} = \begin{bmatrix} 1 & 12 \\ 1 & 11 \\ 1 & 10 \\ 1 & 9 \\ 1 & 8 \\ 1 & 7 \\ 1 & 6 \\ 1 & 5 \\ 1 & 4 \\ 1 & 3 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}$$

donde los vectores de severidades para cada póliza vienen dados por:

$$\mathbf{x}_{1.} = \begin{bmatrix} 1,738 \\ 1,642 \\ 1,794 \\ 2,051 \\ 2,079 \\ 2,234 \\ 2,032 \\ 2,035 \\ 2,115 \\ 2,262 \\ 2,267 \\ 2,517 \end{bmatrix} \quad \mathbf{x}_{2.} = \begin{bmatrix} 1,364 \\ 1,408 \\ 1,597 \\ 1,444 \\ 1,342 \\ 1,675 \\ 1,470 \\ 1,448 \\ 1,464 \\ 1,831 \\ 1,612 \\ 1,471 \end{bmatrix} \quad \mathbf{x}_{3.} = \begin{bmatrix} 1,759 \\ 1,685 \\ 1,479 \\ 1,763 \\ 1,674 \\ 2,103 \\ 1,502 \\ 1,622 \\ 1,828 \\ 2,155 \\ 2,233 \\ 2,059 \end{bmatrix} \quad \mathbf{x}_{4.} = \begin{bmatrix} 1,223 \\ 1,146 \\ 1,010 \\ 1,257 \\ 1,426 \\ 1,532 \\ 1,953 \\ 1,123 \\ 1,343 \\ 1,243 \\ 1,762 \\ 1,306 \end{bmatrix} \quad \mathbf{x}_{5.} = \begin{bmatrix} 1,456 \\ 1,499 \\ 1,609 \\ 1,741 \\ 1,482 \\ 1,572 \\ 1,606 \\ 1,735 \\ 1,607 \\ 1,573 \\ 1,613 \\ 1,690 \end{bmatrix}$$

Con las fórmulas vistas y la especificación (2.5) obtenemos las siguientes estimaciones de los parámetros estructurales:

Varianza dentro de los portafolios: $\hat{s}^2 = 49'870,187$.

Estimaciones condicionalmente insesgados de los $\beta(\Theta_i)$:

$$\mathbf{B}_i = \frac{1}{(s^2)^2} (\mathbb{Y}^\top \mathbb{V}_i^{-1} \mathbb{Y})^{-1} \mathbb{Y}^\top \mathbb{V}_i^{-1} \mathbf{X}_i.,$$

donde \mathbb{V}_i está formada por los inversos de los pesos para la póliza i ; es decir, por:

$$\mathbb{V}_i = \text{diag}\left(\frac{1}{\omega_{i1}}, \frac{1}{\omega_{i2}}, \frac{1}{\omega_{i3}}, \dots, \frac{1}{\omega_{i12}}\right).$$

Hecho esto procedemos a calcular recursivamente los parámetros estructurales \mathbf{b} y \mathbf{A} , pues ellos dependen de la matriz de credibilidad. Estos deben de encontrarse con métodos iterativos que nos brindan estimaciones siguientes.

Prima de riesgo colectivo

$$\hat{\mathbf{b}} = \begin{bmatrix} 1,468.775 \\ 32.049 \end{bmatrix}$$

Varianza del portafolio

$$\hat{\mathbf{A}} = \begin{bmatrix} 24,154.175 & 2,699.975 \\ 2,699.975 & 301.806 \end{bmatrix}$$

Matrices de credibilidad

matrices de cred.	1	2	3	4	5
\mathbb{Z}_i	$\begin{bmatrix} 0.549 & 3.972 \\ 0.061 & 0.444 \end{bmatrix}$	$\begin{bmatrix} 0.531 & 3.912 \\ 0.059 & 0.437 \end{bmatrix}$	$\begin{bmatrix} 0.532 & 3.770 \\ 0.059 & 0.421 \end{bmatrix}$	$\begin{bmatrix} 0.478 & 3.421 \\ 0.053 & 0.382 \end{bmatrix}$	$\begin{bmatrix} 0.539 & 3.965 \\ 0.060 & 0.443 \end{bmatrix}$

Cuadro 2.9: Matrices de credibilidad por el Modelo de Hachemeister

Sobre la base de la proposición 2.1 se obtendrán las siguientes estimaciones lineales de Bayes para los \mathbf{B}_i :

Póliza	1	2	3	4	5
\mathbf{B}_i	$\begin{bmatrix} 1,658.472 \\ 62.392 \end{bmatrix}$	$\begin{bmatrix} 1,398.303 \\ 17.140 \end{bmatrix}$	$\begin{bmatrix} 1,532.999 \\ 43.307 \end{bmatrix}$	$\begin{bmatrix} 1,176.704 \\ 27.807 \end{bmatrix}$	$\begin{bmatrix} 1,521.899 \\ 11.874 \end{bmatrix}$

Cuadro 2.10: Estimaciones encontradas utilizando el modelo de Hachemeister

Con los resultados anteriores y el teorema 2.6 los coeficientes de regresión estructurales para cada una de las pólizas esta dada por:

Póliza	1	2	3	4	5
$\hat{\beta}(\Theta_i)$	$\begin{bmatrix} 1,693.523 \\ 57.171 \end{bmatrix}$	$\begin{bmatrix} 1,373.030 \\ 21.346 \end{bmatrix}$	$\begin{bmatrix} 1,545.364 \\ 40.610 \end{bmatrix}$	$\begin{bmatrix} 1,314.549 \\ 14.809 \end{bmatrix}$	$\begin{bmatrix} 1,417.409 \\ 26.307 \end{bmatrix}$

Cuadro 2.11: Estimaciones encontradas utilizando el modelo de Hachemeister

Finalmente, utilizando la ecuación 2.6 las primas de credibilidad, según el método de Hachemeister, vienen dadas por:

$$\begin{aligned} \hat{\mu}(\Theta_1) &= 2,436.752; \hat{\mu}(\Theta_2) = 1,650.533; \hat{\mu}(\Theta_3) = 2,073.296; \\ \hat{\mu}(\Theta_4) &= 1,507.070; \hat{\mu}(\Theta_5) = 1,759.403. \end{aligned}$$

Capítulo 3

Modelos lineales mixtos

En este capítulo revisaremos brevemente la teoría de los modelos lineales mixtos con la finalidad de mejor entender el enfoque de credibilidad que más adelante desarrollaremos con base en estos modelos. Al igual que en el capítulo anterior utilizaremos para su descripción la perspectiva de los espacios de Hilbert.

3.1. El modelo de regresión lineal

Recordemos que un modelo de regresión lineal especifica que:

$$\mathbf{Y} = \mathbb{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad (3.1)$$

donde \mathbf{Y} es un vector columna $n \times 1$ de las respuestas de n sujetos a una variable dependiente Y , $\mathbb{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_k]$ es una matrix $n \times k$ que contiene a los vectores columnas de orden $n \times 1$, \mathbf{x}_j , que contienen la data de la j -ésima variable independiente o predictora para cada sujeto (incluyendo la posibilidad de que \mathbf{x}_1 sea un vector columna de unos), $\boldsymbol{\beta}$ es un vector $k \times 1$ de coeficientes de regresión desconocidos que miden la contribución de cada variable independiente en el valor medio de Y y $\boldsymbol{\epsilon}$ es un error aleatorio con media 0 de orden $n \times 1$ asociado a cada sujeto. Este último se supone por lo común que tiene una matriz de covarianza $\sigma^2 \mathbb{I}$ y distribución normal; es decir, con componentes independientes.

Si definimos el subespacio de \mathbb{R}^n generado por las columnas de la matrix \mathbb{X} , entonces el teorema de proyección para espacios de Hilbert garantiza la existencia de un único elemento en este espacio que se encuentre lo más cerca, en el sentido medio cuadrático, a \mathbf{Y} . Este es precisamente la proyección ortogonal de este vector en el espacio de las funciones afin-lineales de las variables predictoras al cual denotaremos por $\hat{\mathbf{Y}} = \mathbb{X}\hat{\boldsymbol{\beta}}$. En tal sentido, llamaremos a $\hat{\boldsymbol{\beta}}$ el estimador de mínimos cuadrados de $\boldsymbol{\beta}$.

Del teorema de proyección puede fácilmente deducirse que si $\mathbb{X}^\top \mathbb{X}$ es no singular, entonces

$$\hat{\boldsymbol{\beta}} = (\mathbb{X}^\top \mathbb{X})^{-1} \mathbb{X}^\top \mathbf{Y}.$$

Es conocido que, bajo el supuesto de normalidad en el modelo dado arriba, este estimador coincide con el de máxima verosimilitud. Si bien tal resultado no es cierto en general, es posible, bajo algunas transformaciones, obtener un resultado similar para un modelo heterocedástico que especifique que:

$$\text{Var}(\boldsymbol{\epsilon}) = \Sigma,$$

donde Σ es una matriz simétrica y positiva definida.

En efecto, por el teorema de descomposición espectral podemos garantizar la existencia de una matriz ortogonal \mathbb{P} ($\mathbb{P}^\top \mathbb{P} = \mathbb{P} \mathbb{P}^\top = \mathbb{I}$) y una matriz diagonal Λ tal que $\Sigma = \mathbb{P} \Lambda \mathbb{P}^\top$. Así multiplicando la ecuación de regresión (3.1) por $\Lambda^{-1/2} \mathbb{P}^\top$, tenemos:

$$\Lambda^{-1/2} \mathbb{P}^\top \mathbb{Y} = \Lambda^{-1/2} \mathbb{P}^\top \mathbb{X} \beta + \mathbb{U},$$

donde $\mathbb{U} = \Lambda^{-1/2} \mathbb{P}^\top \epsilon$ puede verificarse que es un término de error de media cero y matriz de covarianza $\sigma^2 \mathbb{I}$.

Como ahora este último modelo satisface los supuestos clásicos uno obtiene el denominado estimador de mínimos cuadrados generalizados:

$$\hat{\beta}_{GLS} = (\mathbb{X}^\top \Sigma^{-1} \mathbb{X})^{-1} \mathbb{X}^\top \Sigma^{-1} \mathbb{Y}. \quad (3.2)$$

Observe que este estimador lineal, de la forma $\hat{\mathbb{C}} \mathbb{Y}$, es insesgado. Más aún, el teorema de Gauss-Markov garantiza que este es el BLUE (de best linear unbiased estimator) de β ; vale decir, el mejor estimador lineal e insesgado de β , donde por mejor entendemos que posee mínima varianza sobre todos los estimadores lineales e insesgados de β . Dicho en otras palabras, la matriz $\hat{\mathbb{C}}$ en $\hat{\beta}_{GLS} = \hat{\mathbb{C}} \mathbb{Y}$ es solución de

$$\hat{\mathbb{C}} = \arg \min_{\{\mathbb{C} / E[\mathbb{C} \mathbb{Y}] = \beta\}} \mathbb{C} \Sigma \mathbb{C}^\top, \quad (3.3)$$

donde se entiende que la minimización va sobre todas las matrices \mathbb{C} de orden $(k+1) \times n$ que hagan de $\mathbb{C} \mathbb{Y}$ un estimador lineal e insesgado de β y que $\mathbb{C} \Sigma \mathbb{C}^\top$ es mínima en el sentido que $\hat{\mathbb{C}} \Sigma \hat{\mathbb{C}}^\top - \mathbb{C} \Sigma \mathbb{C}^\top$ sea una matriz positiva semi-definida para cualquier \mathbb{C} en el conjunto factible anterior.

3.2. El mejor predictor lineal e insesgado de un modelo de regresión

Consideremos un modelo de regresión lineal heterocedástico como el previamente descrito y supongamos estemos interesados en predecir en él un valor particular Y_ω no observado en el modelo para el vector de variables independientes ω . En tal sentido esta variable aleatoria, que asumiremos es de varianza finita, deberá de satisfacer $E[Y_\omega] = \omega^\top \beta$. Una asunción clave para esta predicción, y que simplifica enormemente su viabilidad, es asumir que ella sea una función afín-lineal de las variables dependientes Y_1, Y_2, \dots, Y_n ya observadas; es decir, tenga la forma:

$$Z = a_0 + \sum_{i=1}^n a_i Y_i,$$

donde a_0, a_1, \dots, a_n son constantes por determinar. Más aún, por la consideración anterior deberá cumplirse que

$$E[Z] = a_0 + \sum_{i=1}^n a_i \mu_i = \omega^\top \beta,$$

donde $\mu_i = E[Y_i]$. En otras palabras, estaremos interesados en buscar predictores de la forma:

$$Z = \omega^\top \beta + \sum_{i=1}^n a_i (Y_i - \mu_i). \quad (3.4)$$

Si definimos el subespacio vectorial cerrado \mathcal{M} de L^2 de todas las variables aleatorias Z con la forma (3.4), entonces el teorema de proyección en espacios de Hilbert garantiza la existencia de una única variable aleatoria $\hat{Y}_\omega = \omega^\top \beta + \sum_{i=1}^n \hat{a}_i (Y_i - \mu_i) \in \mathcal{M}$ tal que

$$E[(Y_\omega - \hat{Y}_\omega)Z] = 0, \quad \forall Z \in \mathcal{M}.$$

En particular, para las variables aleatorias $Z = \omega^\top \beta + Y_j - \mu_j \in \mathcal{M}$, con $j = 1, 2, \dots, n$, se cumple que

$$E[(Y_\omega - \omega^\top \beta - \sum_{i=1}^n \hat{a}_i (Y_i - \mu_i))(\omega^\top \beta + Y_j - \mu_j)] = 0$$

o equivalentemente que

$$\sum_{i=1}^n \hat{a}_i \text{Cov}(Y_i, Y_j) = \text{Cov}(Y_\omega, Y_j), \quad \forall j = 1, 2, \dots, n.$$

Este sistema puede representarse matricialmente como

$$\hat{a}\Sigma = \text{Cov}(Y_\omega, \mathbb{Y})^\top$$

y tiene como solución al vector $1 \times n$:

$$\hat{a} = \text{Cov}(Y_\omega, \mathbb{Y})^\top \Sigma^{-1}.$$

Así obtenemos que

$$\hat{Y}_\omega = \omega^\top \beta + \text{Cov}(Y_\omega, \mathbb{Y})\Sigma^{-1}(\mathbb{Y} - \mathbb{X}\beta)$$

es el mejor predictor lineal (en el sentido medio cuadrático) de Y_ω .

Un problema con el predictor anterior es que β es un parámetro desconocido, por lo que surge de manera natural la posibilidad de reemplazarlo por su estimador de mínimos cuadrados generalizados. Si hacemos esto; es decir, consideramos la variable aleatoria

$$\begin{aligned} \hat{Y}_\omega^{BLUP} &= \omega^\top \hat{\beta}_{GLS} + \text{Cov}(Y_\omega, \mathbb{Y})\Sigma^{-1}(\mathbb{Y} - \mathbb{X}\hat{\beta}_{GLS}) \\ &= (\omega^\top - \text{Cov}(Y_\omega, \mathbb{Y})\Sigma^{-1}\mathbb{X})\hat{\beta}_{GLS} + \text{Cov}(Y_\omega, \mathbb{Y})\Sigma^{-1}\mathbb{Y}, \end{aligned}$$

entonces obtenemos el llamado mejor predictor lineal e insesgado (o BLUP) de Y_ω . Una demostración de que esta v.a. efectivamente minimiza la media cuadrática del error al estimar Y_ω puede encontrarse por ejemplo en [Frees et al. \(1999\)](#).

3.3. Modelos Lineales Mixtos

Los modelos lineales mixtos (LMM) son una extensión de los modelos clásicos de regresión lineal estudiados en la sección 3.1, mediante la incorporación de los llamados efectos aleatorios. Estos aparecen al considerar que algunos de los coeficientes en el modelo de regresión lineal no son fijos sino aleatorios debido a que estos representan a ciertos grupos o conglomerados de sujetos interdependientes entre si ya sea porque fueron tomados aleatoriamente de manera jerárquica o corresponden a sujetos relacionados e incluso a los mismos sujetos a lo largo del tiempo. Los modelos lineales mixtos incluyen en su formulación parámetros fijos, comunes a toda la población, y parámetros aleatorios específicos del grupo o conglomerado. Los parámetros aleatorios se consideran realizaciones aleatorias de un proceso de media cero, cuya varianza define la componente del error asociada a cada grupo o conglomerado. En nuestro desarrollo asumiremos independencia entre los distintos grupos o conglomerados por lo que bastará formular el modelo al interior de cualesquiera de estos grupos.

Un modelo lineal mixto sobre n observaciones se escribe como:

$$\mathbf{Y} = \mathbb{X}\boldsymbol{\beta} + \mathbb{Z}\mathbf{b} + \boldsymbol{\epsilon}, \quad (3.5)$$

donde $\mathbf{Y} = [Y_1 \ Y_2 \ \dots \ Y_n]^\top$ es un vector columna $n \times 1$ de variables dependientes, \mathbb{X} es una matriz $n \times k$ que representa los valores de las variables explicativas para los efectos fijos, \mathbb{Z} es una matriz $n \times q$ de variables explicativas para los efectos aleatorios, $\boldsymbol{\beta}$ es un vector columna $k \times 1$ de parámetros fijos, \mathbf{b} es un vector columna $q \times 1$ de parámetros aleatorios y $\boldsymbol{\epsilon}$ es un vector columna $n \times 1$ de errores aleatorios asociados a cada observación dentro del grupo. En nuestro modelo asumiremos que $\mathbf{b} \sim N(0, \mathbb{H})$ y $\boldsymbol{\epsilon} \sim N(0, \Sigma)$ son vectores aleatorios normales independientes.

Se sigue entonces que:

$$\begin{aligned} E[\mathbf{Y}] &= \mathbb{X}\boldsymbol{\beta} \\ E[\mathbf{Y}|\mathbf{b}] &= \mathbb{X}\boldsymbol{\beta} + \mathbb{Z}\mathbf{b} \end{aligned}$$

y

$$Var(\mathbf{Y}) = \mathbb{V} = \mathbb{Z}\mathbb{H}\mathbb{Z}^\top + \Sigma.$$

3.3.1. Estimadores del Modelo

Si bien de las asunciones del modelo $\mathbf{Y} \sim N(\mathbb{X}\boldsymbol{\beta}, \mathbb{V})$, uno podría pensar en utilizar el estimador (3.2) de mínimos cuadrados generalizados para $\boldsymbol{\beta}$, este resulta inadecuado pues \mathbb{V} es en general desconocido y dependiente de un vector de parámetros que denotaremos por $\boldsymbol{\alpha}$. Una solución para este problema es utilizar el estimador de máxima verosimilitud, el cual maximiza la función de verosimilitud de \mathbf{Y} en su valor observado $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n]^\top$:

$$L(\boldsymbol{\beta}, \boldsymbol{\alpha}) = \prod_{i=1}^n (2\pi)^{-\frac{n}{2}} |\mathbb{V}(\boldsymbol{\alpha})|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{y}_i - \mathbb{X}\boldsymbol{\beta})^\top \mathbb{V}^{-1}(\boldsymbol{\alpha}) (\mathbf{y}_i - \mathbb{X}\boldsymbol{\beta}) \right\} \quad (3.6)$$

Esta maximización es en principio directa pero podrían presentarse complicaciones en la práctica, en especial si el modelo incluye un gran número de parámetros para los efectos aleatorios. Un problema particular es encontrar algún parámetro de varianza negativo. Más aún, los estimadores de máxima verosimilitud no son insesgados y el sesgo podría ser significativo. La estimación por máxima verosimilitud restringida (REML) es un intento para solucionar estos problemas. La idea en este método es encontrar una combinación lineal de la componentes de \mathbf{Y} cuya media sea 0 y que por tanto hagan que esta nueva v.a, de también distribución normal, sólo dependa de $\boldsymbol{\alpha}$. Más precisamente sea

$$\mathbf{Q} = \mathbf{I} - \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top$$

y definamos $\mathbf{W} = \mathbf{QY}$. Entonces $\mathbf{W} \sim N(0, \mathbf{QVQ}^\top)$. Más aún se puede probar (Frees et al., 1999) que \mathbf{W} es independiente del (seudo)estimador de mínimos cuadrados generalizados

$$\hat{\boldsymbol{\beta}}_{GLS} = (\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{V}^{-1} \mathbf{Y}.$$

Note además que podríamos usar cualquier otra transformación lineal de la forma \mathbf{AQY} cuya media sea 0. Patterson and Thompson (1971) y Harville (1974, 1977) muestran que el logaritmo de la función de verosimilitud de esta transformación no dependen de \mathbf{A} y viene dada por:

$$\log L_{REML}(\boldsymbol{\alpha}) = \frac{1}{2} (\log |\mathbf{V}(\boldsymbol{\alpha})| + \log |\mathbf{X}^\top \mathbf{V}(\boldsymbol{\alpha}) \mathbf{X}| + (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}_{GLS}(\boldsymbol{\alpha}))^\top \mathbf{V}^{-1}(\boldsymbol{\alpha}) (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}_{GLS}(\boldsymbol{\alpha})).$$

El estimador de máxima verosimilitud restringida de $\boldsymbol{\alpha}$, $\hat{\boldsymbol{\alpha}}_{REML}$, se obtiene entonces al maximizar esta función y el correspondiente estimador de $\boldsymbol{\beta}$ es

$$\hat{\boldsymbol{\beta}}_{REML} = \hat{\boldsymbol{\beta}}_{GLS}(\hat{\boldsymbol{\alpha}}_{REML}) = (\mathbf{X}^\top \mathbf{V}^{-1}(\hat{\boldsymbol{\alpha}}_{REML}) \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{V}^{-1}(\hat{\boldsymbol{\alpha}}_{REML}) \mathbf{Y}.$$

Capítulo 4

Credibilidad bajo modelos lineales mixtos

En esta sección revisaremos nuevamente los modelos de credibilidad introducidos en el capítulo 2, pero presentados ahora desde el enfoque de los modelos lineales mixtos.

En la literatura tradicional sobre credibilidad una asunción muy común es la de suponer que los errores y/o severidades que se observan a lo largo del tiempo en una póliza son variables aleatorias independientes, condicionados al conocimiento del parámetro de heterogeneidad de la póliza. En la práctica, sin embargo, estos términos están usualmente correlacionados dado que la información es tomada de una misma póliza o individuo a lo largo del tiempo. El modelo de efectos mixtos se puede utilizar precisamente para introducir esta correlación dentro de una estructura de datos longitudinales. Por este motivo, y sobre la base de los resultados de Xu (2008), Fung y Xu (2010) y Frees et al. (1999), que en este capítulo mostraremos como los modelos clásicos de credibilidad estudiados; vale decir, los modelos de Bühlmann (1967), Bühlmann y Straub (1970), Jewell (1975) y Hachemeister (1975) pueden verse también como modelos lineales mixtos. Este enfoque no sólo nos permitirá aprovechar las herramientas ya conocidas para el estudio de los modelos lineales mixtos sino también, como más adelante veremos, nos proveerá de una mejora en nuestras estimaciones cuando las asunciones de independencia en el modelo de heterogeneidad pudieran ser violadas.

4.1. El ajuste al modelo de Bühlmann

Dada la presencia en el modelo de Bühlmann de una variable aleatoria estructural Θ_i , que recordemos contiene la información esencial de la póliza i , uno podría interpretar este modelo de forma equivalente como:

$$X_{it} = m + (\mu(\Theta_i) - m) + \epsilon_{it}, \quad t = 1, 2, \dots, n_i \quad (4.1)$$

donde según los supuestos del modelo m es la prima de riesgo colectivo y los errores ϵ_{it} son variables aleatorias no correlacionadas con media 0 y $Var(\epsilon_{it}) = s^2$. De asumirse normalidad en las severidades y en la prima de riesgo individual, el modelo (4.1) representa entonces un modelo de análisis de varianza a una vía (ANOVA). Este es concretamente un modelo de efectos aleatorios desde que $\mu(\Theta_i) - m$ es una variable aleatoria normal de media 0 y varianza a , el cual obviamente pertenece a la familia de los modelos lineales mixtos. Más explícitamente (4.1), precindiendo por comodidad del índice i , se puede expresar en el formato de un modelo lineal mixto por:

$$\mathbf{Y} = \mathbb{X}\boldsymbol{\beta} + \mathbb{Z}\mathbf{b} + \boldsymbol{\epsilon} \quad (4.2)$$

donde $\mathbf{Y} = [X_{i1} \ X_{i2} \ \dots \ X_{in_i}]^\top$ está conformada por las severidades para la póliza i , $\mathbb{X} = \mathbb{Z}$ es un vector columna $n_i \times 1$ de unos, $\beta = m$ es la prima de riesgo colectivo (para todo el portafolio), $\mathbf{b} = (\mu(\Theta_i) - m)$ es el parámetro de efectos aleatorios del modelo y ϵ es un error aleatorio de media 0 y matriz de varianza-covarianza

$$\Sigma = s^2 I_{n_i}.$$

Para la implementación computacional de este modelo y los otros modelos mencionados haremos uso del paquete **lme** de R. En el modelo de Bühlmann, por ejemplo, deberíamos ordenar primero la data de tal manera que esta contenga los datos de todas las severidades observadas, x_{it} , para cada una de las pólizas y una variable indicadora I de la póliza asociada a cada severidad, ambas en columnas. La data que denominaremos **datos1** contendrá entonces a:

$$Y = [x_{11} \ x_{12} \ \dots \ x_{1n_1} \ x_{21} \ x_{22} \ \dots \ x_{2n_2} \ x_{31} \ x_{32} \ \dots \ x_{3n_3} \ \dots \ x_{k1} \ x_{k2} \ \dots \ x_{kn_k}]^\top \quad (4.3)$$

e

$$I = [1 \ 1 \ \dots \ 1 \ 2 \ 2 \ \dots \ 2 \ 3 \ 3 \ \dots \ 3 \ \dots \ k \ k \ \dots \ k]^\top, \quad (4.4)$$

respectivamente.

El ajuste de máxima verosimilitud del modelo (4.1) a **datos1** se puede obtener fácilmente con R a través de la declaración:

```
lme(Y~1,data=datos1, random=list(I=pdDiag( ~1)),method="ML")
```

El primer argumento indica que la respuesta es Y y hay un solo efecto fijo que es el intercepto. El segundo argumento indica que los datos para el estudio se pueden encontrar en el archivo denominado **datos1**. El tercer argumento indica que existe un efecto aleatorio para cada grupo y el grupo está dado por la variable I . Además `pdDiag` especifica que el patrón de la matriz de varianza-covarianza de los efectos aleatorios es una matriz diagonal. Por las características del modelo de Bühlmann, la varianza de los efectos aleatorios puede ser visto como una matriz diagonal 1×1 . El último argumento indica que el modelo se ajusta mediante el método de máxima verosimilitud. Si no se especifica el método de estimación, el modelo se ajusta por defecto por el método de máxima verosimilitud restringida.

Además, si la estructura de error no es independiente, por ejemplo $MA(1)$, podemos usar el objeto `corARMA` referido al modelo autoregresivo de medias móviles para especificar la estructura de correlación con:

```
cs1ARMA <- corARMA(value, p , q )
```

siendo

- **value** : un vector con los valores de los parámetros del modelo autoregresivo y/o media móvil, que deben tener longitud $p+q$ y todos los elementos entre -1 y 1. El valor por defecto es un vector de ceros, lo que corresponde a observaciones no correlacionadas.

- p, q : son números enteros no negativos que especifican el orden autorregresivo y el orden de media móvil de la estructura ARMA, respectivamente. Por defecto son ceros.

Un modelo con una estructura de error de medias móviles de orden 1, MA(1), viene por ejemplo dado por:

```
lme ( Y~1, data =  datos1, random=list(I = pdDiag( ~1)),
      method =  "ML", corr=cs1ARMA )
```

En el estudio nosotros usaremos la estructura

```
cs1ARMA <- corARMA(0.4, p=0 , q=1 )
```

Para especificar la estructura de error, podemos utilizar también otros diferentes objetos que pertenecen a la clase `corStruct` en R. La tabla siguiente es un resumen de los objetos de uso común, que fue tomado de la Tabla 5.3 de [Pinheiro y Bates \(2000\)](#).

Standard corStruct classes	
corSymm	general
corAR1	autoregressive of order 1
corCAR1	continuous-time AR(1)
corARMA	autoregressive-moving average
corExp	exponential
corGaus	Gaussian
corLin	linear

Cuadro 4.1: Tomado de la Tabla 5.3 de [Pinheiro y Bates \(2000\)](#), pag. 234

4.1.1. Aplicación a los datos de Hachemeister

En el apéndice B se presenta el código para estimar mediante modelos lineales mixtos los parámetros estructurales para el modelo de Bühlmann. Utilizando los datos de [Hachemeister \(1975\)](#) que fueron explicados en la sección 2.4.1. Los resultados que encontramos fueron los siguientes:

Estimador	Parámetros			
	m	a	s^2	Z
Clásico	1,671.017	72,310.02	46,040.47	0.9496143
ML	1,671.017	57,080.68	46,040.47	0.9370179
REML	1,671.017	72,310.02	46,040.47	0.9496143

Cuadro 4.2: Estimación de los parámetros estructurales según el modelo de Bühlmann para los datos de Hachemeister.

donde:

- Clásico: se refiere a la estimación de los parámetros estructurales utilizando los estimadores propuestos en la sección 2.4.1 y cuyos cálculos se encuentran en 2.5.2. Estos cálculos pueden también hacerse con el paquete **actuar** de R.

- ML : indica que la estimación fue obtenida a través del método de máxima verosimilitud (ML) bajo el enfoque de los modelos lineales mixtos y el paquete **lme**.
- REML: indica que la estimación fue obtenida a través del método de máxima verosimilitud restringida (REML) bajo el enfoque de los modelos lineales mixtos y el paquete **lme**.

Como se puede observar del cuadro 4.2 los resultados de la estimación de los parámetros utilizando los estimadores clásicos bajo modelos lineales mixtos son exactamente los mismos para todos los estimadores, excepto cuando se aplica el método de máxima verosimilitud, donde se puede observar una menor varianza $a=57,080.68$, en comparación con los otros métodos que presentan un valor de 72,310.02. Esta coincidencia se debe en parte a que el diseño utilizado en esta aplicación es balanceado; es decir, tienen un mismo número de severidades asociadas a cada póliza.

En base a lo anterior se hallaron las siguientes primas:

Estimador	Primas estimadas				
	1	2	3	4	5
Clásico	2,044.04	1,518.59	1,814.23	1,375.99	1,602.23
ML	2,039.09	1,520.61	1,812.34	1,379.90	1,603.15
REML	2,044.04	1,518.59	1,814.23	1,375.99	1,602.23

Cuadro 4.3: Estimación de las primas según el modelo de Bühlmann para los datos de Hachemeister.

4.2. El ajuste al modelo de Bühlmann-Straub

Como vimos el modelo de Bühlmann-Straub es una extensión del modelo de Bühlmann. En este modelo, a cada severidad X_{it} se le asocia una constante o peso w_{it} que mide su exposición. A diferencia del modelo de Bühlmann, la varianza de cada observación es inversamente proporcional a los pesos w_{it} . y simplemente hace que el modelo ANOVA de efectos aleatorios para el modelo Bühlmann sea heterocedástico. Concretamente tenemos que:

$$X_{it} = m + (\mu(\Theta_i) - m) + \epsilon_{it}, \quad t = 1, 2, \dots, n_i \quad (4.5)$$

donde ahora los errores ϵ_{it} siguen siendo variables aleatorias no correlacionadas de media 0 pero ahora con una matriz de varianzas dada por:

$$\Sigma = s^2 \begin{bmatrix} \frac{1}{w_{i1}} & 0 & \dots & 0 \\ 0 & \frac{1}{w_{i2}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{w_{in_i}} \end{bmatrix}$$

Al igual que en el modelo Bühlmann este modelo se puede expresar como un modelo lineal mixto mediante:

$$\mathbf{Y} = \mathbb{X}\boldsymbol{\beta} + \mathbb{Z}\mathbf{b} + \boldsymbol{\epsilon}$$

donde, prescindiendo del índice i , $\mathbf{Y} = [X_{i1} \ X_{i2} \ \dots \ X_{in_i}]^T$, $\mathbb{X} = \mathbb{Z}$ es un vector columna $n_i \times 1$ de unos, $\boldsymbol{\beta} = m$, $\mathbf{b} = (\mu(\Theta_i) - m)$ y $\boldsymbol{\epsilon}$ es el vector aleatorio columna de errores especificados.

En el modelo de Bühlmann-Straub, la variación de cada termino de error ϵ_{it} está asociado con un peso w_{it} . Por lo tanto para su implementación computacional con la rutina **lme** de R, además de las dos variables en el conjunto de datos, **datos1**, que fue utilizado en el modelo de Bühlmann, el conjunto de entradas en el modelo de Bühlmann-Straub deberá contener una variable adicional **W** dada por:

$$\mathbf{W} = \left[\frac{1}{\omega_{11}}, \dots, \frac{1}{\omega_{1n_1}}, \frac{1}{\omega_{21}}, \dots, \frac{1}{\omega_{2n_2}}, \frac{1}{\omega_{k1}}, \dots, \frac{1}{\omega_{kn_k}} \right]^T \quad (4.6)$$

Como lo indica Xu (2008), la inclusión de los pesos w_{it} puede hacerse en R con la opción **varFixed** y la variable:

```
vf1 <- varFixed( ~ W)
```

Así el ajuste de máxima verosimilitud del modelo de Bühlmann-Straub en el paquete **lme** se puede obtener por:

```
lme ( Y ~ 1, data = data1, random=list(I = pdDiag( ~ 1)),  
      method = "ML", weight = vf1 )
```

4.2.1. Aplicación a los datos de Hachemeister

En el Apéndice B se puede observar el código en R para estimar mediante modelos lineales mixtos los parámetros estructurales para el modelo de Bühlmann-Straub. Para ello se utilizó los datos de Hachemeister (1975) que fueron explicados anteriormente, encontrándose los siguientes resultados:

Estimador	Parámetros		
	m	a	s^2
Clásico	1,683.71	89,638.73	139'120,026
ML	1,693.43	50,214.21	139'003,664
REML	1,688.76	64,859.74	139'053,560

Cuadro 4.4: Estimación de los parámetros estructurales según el modelo de Bühlmann-Straub para los datos de Hachemeister.

Al igual que el modelo de Bühlmann, el método ML da una media global mayor y la estimación de las varianzas a y s^2 son menores a los proporcionados por el método tradicional y el método lineal con REML.

La diferencia observada en los parámetros estructurales influyen en la estimación de los factores de credibilidad, el método de máxima verosimilitud proporciona un valor para $Z_4=0.59998$, muy menor a los encontrados con los otros métodos.

En base a lo anterior se pueden hallar los resultados de los cuadros 4.5 y 4.6.

Estimador	Factor de credibilidad para la póliza i				
	Z_1	Z_2	Z_3	Z_4	Z_5
Clásico	0.98474	0.92764	0.89848	0.72791	0.95879
ML	0.97310	0.87785	0.83226	0.59998	0.92880
REML	0.97904	0.90272	0.86498	0.65948	0.94396

Cuadro 4.5: Factores de credibilidad según el modelo de Bühlmann-Straub para los datos de Hachemeister.

Estimador	Primas estimadas				
	1	2	3	4	5
Clásico	2,055.17	1,523.71	1,793.44	1,442.97	1,603.29
ML	2,051.037	1,533.480	1,786.987	1,489.165	1,606.493
REML	2,053.122	1,528.494	1,790.034	1,467.317	1,604.812

Cuadro 4.6: Estimación de las primas según el modelo de Bühlmann-Straub para los datos de Hachemeister.

4.3. El ajuste al modelo jerárquico de Jewell

Como se vio anteriormente el método de Jewell permite clasificar un portafolio en sub-portafolios compuestos por pólizas, los cuales contienen un conjunto de severidades observadas a través del tiempo. Supongamos entonces que un portafolio se ha dividido en P sub-portafolios. La descomposición del riesgo asociado con una severidad dentro de este esquema se obtiene como:

$$X_{pit} = m + (\nu(\Theta_p) - m) + (\mu(\Theta_p, \Theta_{pi}) - \nu(\Theta_p)) + \epsilon_{pit}$$

donde los índices $p = 1, \dots, P$, $i = 1, \dots, k_p$, y $t = 1, \dots, n_{pi}$, denotan respectivamente al portafolio, sub-portafolio y periodo de observación de la severidad, m es la media general, $\nu(\Theta_p)$ la prima de riesgo del sub-portafolio p y $\mu(\Theta_p, \Theta_{pi})$ la prima de riesgo individual de la póliza i al interior del sub-portafolio p . Estas últimas componentes se denotan como desviaciones de las medias y por tanto tienen un valor esperado igual a cero, siendo independientes.

Este modelo puede expresarse para cualquier sub-portafolio p y póliza i como un modelo lineal mixto, mediante:

$$\mathbf{Y} = \mathbb{X}\boldsymbol{\beta} + \mathbb{Z}\mathbf{b} + \boldsymbol{\epsilon}$$

donde $\mathbf{Y} = [X_{pi1} \ X_{pi2} \ \dots \ X_{pin_{pi}}]^\top$, $\mathbb{X} = \mathbf{1}_{n_{pi} \times 1}$ es un vector columna de unos $\boldsymbol{\beta} = m$, \mathbb{Z} es una matriz $n_{pi} \times 2$ de unos, $\mathbf{b} = [\nu(\Theta_p) - m, \mu(\Theta_p, \Theta_{pi}) - \nu(\Theta_p)]^\top$ es el vector columna de los parámetros de efectos aleatorios independientes y $\boldsymbol{\epsilon}$ es un error aleatorio de media 0 y matriz de varianza-covarianza:

$$\Sigma = s^2 \begin{bmatrix} \frac{1}{\omega_{pi1}} & 0 & \dots & 0 \\ 0 & \frac{1}{\omega_{pi2}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\omega_{pin_{pi}}} \end{bmatrix}$$

Similarmente al modelo de Bühlmann-Straub, los ω_{pit} son los pesos asignados a cada observación X_{pit} .

Con relación a la implementación computacional en R para este modelo deberemos considerar de entrada como antes a las variables **Y** y **W** en **datos1**. La variable **Y** contendrá todas las severidades observadas, mientras que la variable **W** a sus pesos respectivos. Sin embargo, dada la estructura jerárquica del modelo estas deberían de especificarse como siguen:

$$\mathbf{Y} = [x_{111}, x_{112}, \dots, x_{11n_{11}}, x_{121}, x_{122}, \dots, x_{12n_{12}}, \dots, x_{Pk_P n_{Pk_P}}]^\top \quad (4.7)$$

y

$$\mathbf{W} = \left[\frac{1}{\omega_{111}}, \frac{1}{\omega_{112}}, \dots, \frac{1}{\omega_{11n_{11}}}, \frac{1}{\omega_{121}}, \frac{1}{\omega_{122}}, \dots, \frac{1}{\omega_{12n_{12}}}, \dots, \frac{1}{\omega_{Pk_P n_{Pk_P}}} \right]^\top. \quad (4.8)$$

Además para el enfoque mediante modelos lineales mixtos hay que añadir a **datos1** dos variables **I** y **J**. La variable **I** nos indicará los sub-portafolios a los cuales pertenece las severidades, vale decir:

$$\mathbf{I} = [1_{1 \times k_1}, 2 \times 1_{1 \times k_2}, \dots, P \times 1_{1 \times k_p}]^\top \quad (4.9)$$

Dentro de cada sub-portafolio se utilizará además la variable **J** para indicar las pólizas a las que pertenece las severidades, es decir:

$$\mathbf{J} = \begin{bmatrix} 1_{n_{11} \times 1}, 2 \times 1_{n_{12} \times 1}, \dots, k_1 \times 1_{n_{1k_1} \times 1}, \\ 1_{n_{21} \times 1}, 2 \times 1_{n_{22} \times 1}, \dots, k_2 \times 1_{n_{2k_2} \times 1}, \\ \vdots \\ 1_{n_{P1} \times 1}, 2 \times 1_{n_{P2} \times 1}, \dots, k_p \times 1_{n_{Pk_p} \times 1} \end{bmatrix}^\top$$

Para ajustar el modelo jerárquico, los efectos fijos y aleatorios se expresan como:

$$random = 1|J/I$$

Este argumento indica que sólo hay un efecto fijo, y la estructura de los datos es jerárquica con **J** (que nos indica a que póliza pertenece cada observación) anidado en **I** (que nos indica los sub-portafolios). Las especificaciones de los otros argumentos son similares a los vistos en los modelos de Bühlmann y Bühlmann-Straub. Así tenemos que el ajuste de máxima verosimilitud del modelo jerárquico esta dado por:

$$lme(Y \sim 1, data=datos1, random=\sim 1|J/I, method = "ML", weight = vf1),$$

donde

$$vf1 <- varFixed(\sim W).$$

4.3.1. Aplicación a los datos de Hachemeister

En el apéndice B se presenta el código para la estimación mediante modelos lineales mixtos de los parámetros estructurales para el modelo de Jewell, utilizando los datos de

Hachemeister (1975) explicados en la sección 2.4.1. Empleando estos códigos se obtuvo los siguientes resultados:

Estimador	Parámetros			
	m	a	b	s^2
Clásico	1,742.22	87,263.7	13,414.84	139'120,026
ML	1,693.433	25,107.26	25,106.94	139'003,665
REML	1,688.756	32,430.09	32,430.09	139'053,501

Cuadro 4.7: Estimación de los parámetros estructurales según el modelo de Jewell para los datos de Hachemeister.

Con lo que se obtuvieron las siguientes primas por subportafolio

Estimador	Primas estimadas				
	1	2	3	4	5
Clásico	2,048.324	1,523.800	1,874.625	1,496.563	1,585.169
ML	2,051.037	1,533.480	1,786.987	1,489.165	1,606.493
REML	2,053.122	1,528.494	1,790.034	1,467.317	1,604.812

Cuadro 4.8: Estimación de las primas según el modelo de Jewell para los datos de Hachemeister.

4.4. El modelo de regresión de Hachemeister

El modelo Hachemeister vimos que tiene la forma:

$$E[\mathbf{X}_i | \Theta_i] = \mathbb{Y} \beta_i, \quad (4.10)$$

donde \mathbf{X}_i es el vector de severidades de la póliza i , Θ_i denota la característica de riesgo no observable para esta póliza, \mathbb{Y} es una matriz de variables explicativas y $\beta_i = \beta(\Theta_i)$ es un vector de coeficientes estructurales de regresión de la póliza. Al igual que los modelos anteriores este modelo se puede expresar como un modelo lineal mixto a escribirse, prescindiéndose del subíndice i , por:

$$\mathbf{Y} = \mathbb{X} \boldsymbol{\beta} + \mathbb{Z} \mathbf{b} + \boldsymbol{\epsilon}$$

donde $\mathbf{Y} = \mathbf{X}_i = [X_{i1} \ X_{i2} \ \dots \ X_{ini}]^\top$, $\mathbb{X} = \mathbf{0}$, $\mathbb{Z} = \mathbb{Y}$, $\mathbf{b} = \beta_i$ y $\boldsymbol{\epsilon}$ es un vector de errores aleatorios de media cero y matriz de varianzas-covarianzas:

$$\Sigma = s^2 \begin{bmatrix} \frac{1}{\omega_{i1}} & 0 & \dots & 0 \\ 0 & \frac{1}{\omega_{i2}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\omega_{ini}} \end{bmatrix}$$

El ajuste por máxima verosimilitud del modelo Hachemeister puede realizarse con R nuevamente mediante el uso del paquete lme escribiendo:

```
lme ( Y ~ -1+X, data = datos1, random=list(I = pdSymm(~ -1+X)),
      method = "ML", weight = vf1)
```

donde

```
vf1 <- varFixed(~ W)
```

 \mathbf{y}

$$\mathbf{W} = [\frac{1}{\omega_{11}}, \frac{1}{\omega_{12}}, \dots, \frac{1}{\omega_{1n_1}}, \frac{1}{\omega_{21}}, \frac{1}{\omega_{22}}, \dots, \frac{1}{\omega_{2n_2}}, \dots, \frac{1}{\omega_{kn_k}}]^\top. \quad (4.11)$$

representa a la matriz diagonal de pesos para cada una de las k pólizas en estudio.

En los argumentos primero y tercero, \sim se puede leer como "se modela como". El número 1 en los argumentos primero y tercero significa que no hay término de intersección en el modelo, de lo contrario la intersección será asumida en el modelo de forma pre-determinada, **datos1** al igual que en los modelos anteriores contiene las severidades y los pesos. El tercer argumento `pdSymm` se utiliza para especificar que la matriz de varianzas-covarianzas de los efectos aleatorios es una matriz simétrica definida positiva.

4.4.1. Aplicación a los datos de Hachemeister

En el apéndice B se presenta el código para la estimación mediante modelos lineales mixtos de los parámetros estructurales por el método de regresión de Hachemeister, utilizando los datos de [Hachemeister \(1975\)](#) explicados en la sección 2.5.1. Empleando estos códigos se obtuvo los siguientes resultados:

Estimador	Parámetros		
	β_0	β_1	s^2
Clásico	1,468.77	32.049	49'870,187
ML	1,502.115	27.88	46'733,936.7
REML	1,501.296	27.75	47'599,071.86

Cuadro 4.9: Estimación según el modelo de Hachemeister

Estimador	Primas estimadas				
	1	2	3	4	5
Clásico	2,436.752	1,650.533	2,073.296	1,507.070	1,759.403
ML	2,463.102	1,607.707	2,065.226	1,466.606	1,720.181
REML	2,463.917	1,605.918	2,067.336	1,453.528	1,719.698

Cuadro 4.10: Estimación según el modelo de Hachemeister

Se puede concluir que los parámetros estimados con los diferentes métodos son diferentes, en el capítulo 5 se realiza simulaciones para analizar cómo estos métodos recuperan los parámetros estructurales y así identificar mejor las ventajas de estos métodos en términos de sus errores cuadráticos medios (MSE).

Capítulo 5

Estudio de Simulación

5.1. Preliminares

En este capítulo realizaremos un estudio de simulación de recuperación de parámetros con los modelos de credibilidad estudiados. La finalidad de este estudio es doble. En primer lugar nos permitirá apreciar cómo funcionan las estimaciones con el enfoque bajo modelos lineales mixtos en comparación con las estimaciones clásicas anteriormente discutidas y en segundo lugar nos permitirá estudiar la eficiencia de este enfoque cuando el supuesto de independencia condicional en el modelo de heterogeneidad no es satisfecha. Vale comentar que de no cumplirse la última hipótesis, todos los estimadores anteriormente desarrollados no necesariamente son óptimos y por tanto los estimadores clásicos propuestos podrían ser ineficientes. En tal sentido, este estudio tratará de mostrar que bajo el enfoque de los modelos lineales mixtos es posible mejorar las estimaciones clásicas anteriores de tomarse en cuenta la violación al supuesto de independencia condicional.

Para poder lograr los objetivos anteriormente planteados, consideraremos en este estudio dos escenarios: uno en el que asumiremos a las severidades como variables aleatorias condicionalmente i.i.d y otra en la que estos serán generados por un proceso de media móviles. Concretamente de acuerdo a estos escenarios generaremos 500 réplicas de un portafolio con 50 pólizas y 5 observaciones de severidades por póliza, obteniendo luego la media y el error cuadrático medio (MSE) de las estimaciones de los parámetros estructurales.

Es importante mencionar que nuestro estudio es muy similar al realizado por Xu (2008), aún cuando este autor manifieste utilizar severidades bajo una estructura de medias móviles, en realidad utiliza otro proceso dependiente por una diferente especificación de la matriz de covarianzas. En nuestro caso nosotros emplearemos una estructura de medias móviles en el sentido estricto de la palabra.

5.2. Estudio de simulación con el modelo de Bühlmann

Con el fin de realizar el estudio de simulación para el modelo de Bühlmann, asumiremos que las severidades bajo este modelo, tal como lo discutimos en el capítulo anterior, pueden ser representados mediante un modelo de varianza a una vía de efectos aleatorios, el cual es descrito como:

$$X_{it} = m + (\Theta_i - m) + \epsilon_{it}, \quad t = 1, 2, 3, 4, 5 \quad (5.1)$$

donde $\epsilon_{it} \sim N(0, s^2)$, $(\Theta_i - m) \sim N(0, a)$ y X_{it} es la t -ésima severidad para la póliza i .

Para el estudio utilizaremos el siguiente juego de parámetros estructurales tomados de Xu (2008)

$$m = 400, a = 25 \text{ y } s^2 = 20.$$

Así como un factor de credibilidad de $Z = \frac{25}{29}$.

La idea central en este capítulo es simular las severidades X_{it} sobre la base de los parámetros estructurales propuestos anteriormente, las características de los métodos propuestos se verificarán en términos de las medias de los estimadores y los errores cuadráticos medios (MSE) encontrados en el proceso de estimación. Se considerarán aquí métodos:

Bühlmann: Se refiere al modelo de Bühlmann y los estimadores clásicos de los parámetros estructurales propuestos en el capítulo 2. Para el cálculo de estos estimadores se utilizó la librería actuar que se encuentra implementada en el programa R.

Lineal mixto: Se refiere al ajuste mediante modelos lineales mixtos que fue visto en el capítulo 4. Para el cálculo de los estimadores se utilizó en un primer escenario el programa R a través de la declaración siguiente:

```
lme(Y~ 1, data=datos1, random=list(I = pdDiag( ~1)), method = "ML")
```

donde el primer argumento indica que la respuesta es Y (el vector de severidades observadas) y hay un solo efecto fijo, el intercepto. El segundo argumento indica que los datos se pueden encontrar en el conjunto de datos denominado **datos1**. El tercer argumento indica que existe un efecto aleatorio para cada grupo y el grupo está dado por la variable I . Además `pdDiag` especifica que el patrón de la matriz 1×1 de varianza-covarianza de los efectos aleatorios es diagonal. El último argumento indica que el modelo se ajusta mediante el método de máxima verosimilitud (ML). Si no se especifica el método de estimación, el modelo se ajusta por defecto al de máxima verosimilitud restringida (REML).

En un segundo escenario se utiliza el modelo denominado de medias móviles, que explica el comportamiento de los errores en términos de una sucesión correspondiente a períodos precedentes, ponderados convenientemente (en este caso se utiliza $\rho = 0.4$). Estos modelos se denotan normalmente con las siglas MA, seguidos del orden entre paréntesis. Así, un modelo de primer orden se representa como MA(1) y responde a la siguiente expresión:

$$\epsilon_{it} = a_{it} + \rho a_{it-1} \quad (5.2)$$

Este modelo en el lenguaje R se puede incorporar especificando la siguiente estructura:

```
eAR <- corARMA(value, p = 0, q = 1)
```

y luego invocando el comando

```
lme(Y~1, data=datos1, random=list(I = pdDiag(~1)), method = "ML", corr = eAR).
```


Los estimadores resultantes en cada muestra se utilizan para calcular el factor de credibilidad Z , lo que nos permite finalmente calcular el estimador de credibilidad empírico lineal de Bayes. Los estimadores encontrados se compararán luego con sus valores reales en términos de sus medias y de sus errores cuadráticos medios (MSE) calculados en base a los verdaderos valores de los parámetros.

En el Apéndice B se puede encontrar el programa en R utilizado para este estudio de simulación, con él se obtuvieron los siguientes resultados cuando las severidades se simularon bajo una estructura de errores independientes:

Bühlmann	Parámetros	m	Z	a	s^2
	valor real	400	25/29	25	20
Clásico	Medias	400.0621	0.8585	25.4313	20.0235
	MSE	3.7377	0.0010	35.8274	4.1062
ML	Medias	400.0621	0.8556	24.8426	20.0235
	MSE	3.8378	0.0011	34.2542	4.1062
ML-MA	Medias	400.0621	0.8554	24.8412	20.0242
	MSE	4.0587	0.0012	35.1294	4.5365
REML	Medias	400.0621	0.8585	25.4313	20.0235
	MSE	3.7377	0.0010	35.8274	4.1062
REML-MA	Medias	400.0621	0.8583	25.4302	20.0241
	MSE	3.9433	0.0011	36.7071	4.5361

Cuadro 5.1: Estimación según el modelo de Bühlmann con errores independientes

donde:

- Clásico : se refiere a la estimación de los parámetros estructurales utilizando los estimadores clásicos de Bühlmann mostrados en la sección 2.4.1.
- ML : indica que la estimación fue obtenida con el paquete lme a través del método de máxima verosimilitud (ML) bajo el enfoque de los modelos lineales mixtos.
- ML-MA : indica que la estimación fue obtenida a través del método de máxima verosimilitud (ML) bajo el enfoque de los modelos lineales mixtos, asumiendo para los errores una estructura de medias móviles con $\rho = 0.4$.
- REML: indica que la estimación fue obtenida a través del método de máxima verosimilitud restringida (REML) bajo el enfoque de los modelos lineales mixtos.
- REML-MA: indica que la estimación fue obtenida a través del método de máxima verosimilitud restringida (REML) bajo el enfoque de los modelos lineales mixtos, asumiendo para los errores una estructura de medias móviles con $\rho = 0.4$.

Con referencia al cuadro 5.1, cuando los términos de error son simulados mediante una estructura de errores independientes, los métodos ML, REML y el enfoque clásico de Bühlmann se observa que funcionan igualmente bien en la estimación de a , Z , y S^2 . Estos

resultados empíricos muestran que el método clásico no siempre tienen el mejor rendimiento, incluso cuando su supuesto de independencia coincide con la estructura de error real. Podemos ver que los otros métodos también muestran un rendimiento competitivo.

Si las severidades se simulan bajo una estructura MA(1) los resultados que se encuentran son:

Bühlmann	Parámetros	m	Z	a	s^2
	valor real	400	25/29	25	20
Clásico	Medias	399.9585	0.8890	28.0790	16.7427
	MSE	6.0259	0.0014	48.0980	13.8587
ML	Medias	399.9590	0.8867	27.4504	16.7427
	MSE	5.6995	0.0014	43.1011	13.8588
ML-MA	Medias	399.9585	0.8569	24.6567	19.4287
	MSE	5.1642	0.0016	37.2903	5.5437
REML	Medias	399.9590	0.8890	28.0790	16.7427
	MSE	6.0259	0.0014	48.0980	13.8587
REML-MA	Medias	399.9585	0.8600	25.2808	19.4281
	MSE	5.0164	0.0015	38.7508	5.543

Cuadro 5.2: Estimación según el modelo de Bühlmann con la estructura de error MA(1) y $\rho = 0.4$

donde:

- Clásico : se refiere a la estimación de los parámetros estructurales utilizando los estimadores clásico de Bühlmann que fueron propuestos en la sección 2.4.1.
- ML : indica que la estimación fue obtenida a través del método de máxima verosimilitud (ML) bajo el enfoque de los modelos lineales mixtos.
- ML-MA : indica que la estimación fue obtenida a través del método de máxima verosimilitud (ML) bajo el enfoque de los modelos lineales mixtos, asumiendo una estructura para los errores de medias móviles con $\rho = 0.4$ y el paquete **lme**.
- REML: indica que la estimación fue obtenida a través del método de máxima verosimilitud restringida (REML) bajo el enfoque de los modelos lineales mixtos.
- REML-MA: indica que la estimación fue obtenida a través del método de máxima verosimilitud restringida (REML) bajo el enfoque de los modelos lineales mixtos, asumiendo una estructura para los errores de medias móviles con $\rho = 0.4$ y el paquete **lme**.

Cuando la hipótesis de errores independientes no es respetada, el cuadro 5.2, muestra que el método clásico de Bühlmann no puede competir con los métodos lineales, en especial, en lo que se refiere a los valores de MSE de los parámetros a y s^2 . En efecto los MSE de los métodos ML-MA y REML-MA son mucho menores. Con $\rho = 0.4$, ML-MA y REML-MA superan a los otros métodos, debido a sus supuestos correctos acerca de la estructura de error. Se puede observar que solo en la estimación con ML-MA y REML-MA los valores de MSE son menores a los encontrados con el método clásico.

5.3. Estudio de simulación con el modelo de Bühlmann-Straub

Como ya se menciona el modelo de Bühlmann-Straub es una extensión del modelo de Bühlmann, es por este motivo que hemos considerado en este punto los mismos parámetros que fueron revisados en el modelo de Bühlmann, excepto que ahora se considera para cada severidad un peso. Estos pesos han sido simulados mediante una distribución uniforme entre 5 y 100. Las simulaciones se realizan bajo dos escenarios, la primera bajo una estructura de errores independientes y la segunda bajo una estructura de errores MA(1) con $\rho = 0.4$.

Así tenemos que los parámetros considerados en este caso son

$$m = 400, a = 25 \text{ y } s^2 = 20.$$

Para analizar la eficacia de los métodos propuestos en la estimación de los factores de credibilidad y las primas correspondientes se utilizó el sesgo de la estimación (en términos de promedio y MSE con respecto a cero)

En el Apéndice B se puede encontrar las rutinas en R que han sido utilizados para este estudio de simulación. En primer lugar se ha considerado para la simulación una estructura de errores independientes, donde se encontraron los siguientes resultados:

Bühlmann-Straub	Parámetros valor real	m	a	s^2
		400	25	20
Clásico	Medias	399.9730	24.4901	20.0307
	MSE	5.0290e-01	25.6478	3.7469
ML	Medias	399.9730	24.0507	20.0307
	MSE	5.0290-01	24.9457	3.7467
ML-MA	Medias	399.9729	24.0495	20.0332
	MSE	5.0300-01	24.9287	4.0134
REML	Medias	399.9730	24.5432	20.0307
	MSE	5.0290e-01	25.2447	3.7467
REML-MA	Medias	399.97295	24.5420	20.0332
	MSE	5.0299e-01	25.2261	4.0133

Cuadro 5.3: Estimación según el modelo de Bühlmann-Straub con errores independientes

Con referencia al cuadro 5.3, podemos observar que cuando el objetivo es estimar el parámetro a , todos los métodos presentan menor MSE que el método clásico, destacando los métodos ML y ML-MA como los mejores en términos de MSE, así también en la estimación de s^2 se puede observar que los mejores métodos de estimación en términos de MSE son los métodos clásico, ML y REML. Los métodos mixtos son equivalentes al clásico e incluso ligeramente superiores en algunos de los estimadores.

Por otro lado, en el Apéndice B se puede encontrar las rutinas en R que han sido utilizados para el estudio de simulación considerando ahora estructura de errores de medias móviles MA(1) con $\rho = 0.4$. Se encontraron los siguientes resultados:

Bühlmann- Straub	Parámetros	m	a	s^2
	valor real	400	25	20
Clásico	Medias	400.0020	24.76425	17.46906
	MSE	5.682264e-01	26.63568	9.690963
ML	Medias	400.0020	24.27207	17.46909
	MSE	5.682278e-01	25.09666	9.690975
ML-MA	Medias	400.0025	24.22685	20.00044
	MSE	5.677151e-01	25.18938	5.090915
REML	Medias	400.0020	24.76887	17.46909
	MSE	5.682288e-01	25.63314	9.690973
REML-MA	Medias	400.0025	25.0923	20.00046
	MSE	5.677162e-01	25.68191	5.090720

Cuadro 5.4: Estimación según el modelo de Bühlmann-Straub con la errores con una estructura MA(1) y $\rho = 0.4$

Como podemos ver en el cuadro 5.4, en lo que se refiere a la estimación del parámetro a se puede observar que los métodos ML y ML-MA producen valores más pequeños de MSE, la diferencia entre los métodos analizados se puede observar en la estimación de s^2 se puede observar que los mejores métodos de estimación en términos de MSE son los métodos ML-MA y REML-MA. Los resultados muestran que el método clásico Bühlmann-Straub no puede competir con los otros métodos de estimación en especial cuando en la estimación se considera la estructura de errores MA(1), resultados que coinciden con los encontrados por (Xu, 2008).

5.4. Estudio de simulación con el modelo de Jewell

La simulación de los datos en este caso siguen la siguiente estructura:

- El portafolio está formada por 50 sectores, es decir, $p = 50$;
- Dentro de cada sector, hay cinco contratos para todo i , es decir, $k_p = 5$.
- Hay 5 observaciones para cada contrato, es decir, $n_{p_i} = 5$.
- La media general es 30, es decir, $m = 30$;
- Los $(\nu(\Theta_p) - m)$ son generados a partir de una distribución normal con media 0 y varianza 25, es decir, $b = 25$;
- Los $(\mu(\Theta_p, \Theta_{pi}) - \nu(\Theta_p))$ son generados a partir de una distribución normal multivariante con media $0_{5 \times 1}$ y matriz de covarianzas $\mathbb{I}_{5 \times 5}$, es decir, $a = 1$,
- Los ϵ_{pit} son simuladas de la distribución normal multivariada con media $0_{5 \times 1}$, matriz de varianzas, Σ , donde σ^2 toma el valor de 16.

Simulamos los datos en dos escenarios diferentes, primero con $\rho = 0$ y luego con $\rho = 0.4$. Cada elemento de ponderación w_{ijt} se genera a partir de una Poisson con media λ_{ij} la cual sigue una distribución uniforme definida en el intervalo (5, 100) y es diferente para cada (i, j) ;

El escenario anterior es simulado 500 veces.

Los métodos utilizados para la estimación de los parámetros a , b , y σ^2 son similares a los de Bühlmann y Bühlmann-Straub. Después de obtener la estimación de estos parámetros, la bondad de esta estimación se muestra en términos del error cuadrático medio (MSE).

En el Apéndice B se puede encontrar las rutinas en R que han sido utilizados para este estudio de simulación. Considerando una estructura de errores independientes, encontrando los siguientes resultados:

Parámetros		<i>ML</i>	<i>ML – MA1</i>	<i>REML</i>	<i>REML – MA1</i>	Clásica
m=30	Media	29.9196964	29.9180323	29.9196964	29.9180323	29.91967
	MSE	4.674227e-01	4.676581e-1	4.674228e-01	4.676582e-1	4.674276e-01
a=1	Media	1.0013434	1.0186960	1.0013438	1.0186967	0.9944415
	MSE	1.229569e-02	1.265425e-02	1.229559e-02	1.265435e-02	1.430485e-02
b=25	Media	24.6112380	24.4268	25.1180291	25.1192180	25.12462
	MSE	26.26010	2.3244e+01	27.19948	27.20819	27.27680
$s^2 = 16$	Media	17.0609549	16.4445578	17.0609549	16.4445578	17.0541200
	MSE	1.134771	0.2019865	1.134771	0.2019818	1.119880
Sesgo de prima	Media	0.0106953	0.0025300	0.0106951	0.0025304	0.0125544
	MSE	4.642345e-03	4.391833e-03	4.642300e-03	4.391810e-03	4.800619e-03
Sesgo prima subportafolio	Media	0.0010431652	0.0013524	0.0005235	0.0008266	5.257265e-04
	MSE	8.043261e-05	8.228346e-05	7.763599e-05	7.905268e-05	3.533204e-05

Cuadro 5.5: Simulación de los datos según modelo de Jewell con errores independientes

El cuadro 5.5 muestra los resultados cuando se considera una estructura de errores independientes, así tenemos: en términos de la MSE, que en la estimación de los parámetros estructurales a , b y S^2 el método ML presenta valores de MSE menores a los encontrados con el método clásico. En el caso del método REML se puede observar que los valores de MSE vistos y el método clásico son aproximadamente iguales; sin embargo, existen algunos casos donde los valores de MSE para el método REML son ligeramente menores, por lo tanto se puede concluir que los resultados en la estimación con el método clásico Jewell pueden ser superados por los otros estimadores.

Por otro lado, en el Apéndice B se puede encontrar las rutinas en R que han sido utilizados para el estudio de simulación, considerando una estructura de errores de medias móviles MA(1) y $\rho = 0.4$. Los siguientes son los resultados encontrados:

Parámetros		<i>ML</i>	<i>ML – MA1</i>	<i>REML</i>	<i>REML – MA1</i>	Clásica
m=30	Media	30.047128	30.046894	30.047128	30.046893	30.047132
	MSE	5.056532e-01	5.055792e-01	5.056527e-01	5.055785e-01	5.056490e-01
a=1	Media	1.0589809	0.9916036	1.0589805	0.9916036	1.0471349
	MSE	1.655208e-02	1.296118e-02	1.655207e-02	1.296119e-02	1.679458e-02
b=25	Media	24.6047448	24.6098470	25.1114843	25.1166237	25.1089630
	MSE	23.09807	23.10615	23.90119	23.91388	23.84160
$s^2 = 16$	Media	13.4861440	16.0162280	13.4861458	16.0162248	13.4787743
	MSE	6.743793	0.6961001	6.743787	0.6960898	6.780446
Sesgo de prima	Media	0.0303118	0.0077071	0.0303118	0.0077069	0.0286527
	MSE	4.835578e-03	3.872597e-03	4.835568e-03	3.872558e-03	4.810412e-03
Sesgo prima subportafolio	Media	0.0022402	0.0012567	0.0018290	0.0008651	0.0001425
	MSE	6.960867e-05	6.344992e-05	6.660890e-05	6.164429e-05	3.759366e-05

Cuadro 5.6: Simulación de los datos según modelo de Jewell, MA(1) y $\rho = 0.4$

El cuadro 5.6 muestra los resultados cuando los términos de error siguen una estructura MA(1) con $\rho = 0.4$. Como podemos ver, en este caso, los valores de MSE en el caso de los parámetros a , b y $s^2 = 16$ son menores cuando en método de estimación es ML-MA, resultados similares se pueden encontrar cuando la estimación se realiza con el método REML-MA, la mayor diferencia en términos de MSE se observa en la estimación de $s^2 = 16$. Por lo tanto se puede concluir que los resultados en la estimación con el método clásico Jewell pueden ser superados por los otros métodos en especial cuando en la estimación se considera la estructura de errores MA(1).

5.5. Estudio de simulación con el modelo de Regresión de Hachemeister

Para la simulación de los datos en este caso adoptamos el modelo de tendencia lineal propuesto por [Hachemeister \(1975\)](#), el cual tiene la forma:

$$X_{it} = b_{i1} + b_{i2}t + \epsilon_{it} \quad (5.3)$$

En esta simulación:

- El número de pólizas es igual a j
- El número de observaciones por contrato es igual a 12 para todo i .
- $b = [1400, 150]^T$
- La varianza comun es $\sigma^2 = 300^2$
- Las varianzas de los coeficientes de b son: $\alpha_1 = 100^2$, $\alpha_{12} = 0$ y $\alpha_2 = 20^2$ y por tanto:

$$\begin{aligned} \mathbb{A} &= \text{Var}(b) \\ &= \begin{bmatrix} \alpha_1 & \alpha_{12} \\ \alpha_{12} & \alpha_2 \end{bmatrix} \\ &= \begin{bmatrix} 100^2 & 0 \\ 0 & 20^2 \end{bmatrix} \end{aligned}$$

El vector de coeficientes de regresión es simulado desde una distribución normal multivariante con media b y varianza \mathbb{A} . Los pesos ω_{ij} son tomados de la data de [Hachemeister \(1975\)](#). El vector de errores es generado a partir estructura de error independiente cuyos resultados se muestran en el cuadro 5.7, y en un segundo caso como una estructura de errores MA(1) con ($\rho = 0.4$). Los resultados se muestran en los cuadros 5.7 y 5.8. El número de simulaciones es de 500.

A continuación se muestra los resultados de la simulación de datos con estructura independiente ($\rho = 0$)

Para		<i>ML</i>	<i>ML – MA</i>	<i>REML</i>	<i>REML – MA</i>	Clásica
$\beta_0=1400$	Media	1398.513	1398.514	1398.553	1398.527	1398.514
	MSE	1.7835e+03	1.7850e+03	1.7815e+03	1.7850e+03	1.7852e+03
$\beta_1=150$	Media	149.6211	149.6210	149.6167	149.6198	149.6210
	MSE	9.4001e+01	9.3962e+01	9.4124e+01	9.4032e+01	9.395191e+01
$s^2 = 90000$	Media	2.3743e+05	2.3732e+05	2.3780e+05	2.3743e+05	2.3726e+05
	MSE	2.9132e+10	2.9084e+10	2.9837e+10	2.9631e+10	2.9056e+10
$\alpha_1 = 10000$	Media	8178.928	10248.50	8175.346	10246.38	10256.26
	MSE	3.6985e+07	5.2715e+07	3.6993e+07	5.2694e+07	5.2815e+07
$\alpha_2 = 400$	Media	312.0980	390.5093	312.0858	390.4366	390.5251
	MSE	5.4881e+04	7.3734e+04	5.4900e+04	7.3730e+04	7.3683e+04
$\alpha_{12} = 0$	Media	2.7460e-02	2.4622e-02	2.7954e-02	2.4802e-02	1.031286
	MSE	2.5140e-01	2.4880e-01	2.5173e-01	2.4861e-01	9.4239e+01

Cuadro 5.7: Simulación de los datos según modelo de Hachemeister con errores independientes

En el cuadro 5.7 se muestra los resultados de la simulación bajo un modelo de Regresión de Hachemeister y estructura de errores independientes, así tenemos: en la estimación de los coeficientes del modelo se puede observar que los valores de MSE de los métodos ML y ML-MA son menores en términos de MSE a los encontrados con el método clásico; en la estimación de s^2 se puede observar que los valores de MSE son aproximadamente iguales; en la estimación de los coeficientes de la matriz \mathbf{A} se observa que en todos los casos los valores de MSE para el método ML son menores a los encontrados en el método clásico resultados similares se encuentran cuando se aplica el método REML. Como se puede observar el método clásico no resulta ser el mejor en la estimación de los parámetros

A continuación se muestra los resultados de la simulación de datos con estructura de errores MA(1) y cuando $\rho = 0.4$

Parámetros		<i>ML</i>	<i>ML – MA</i>	<i>REML</i>	<i>REML – MA</i>	Clásica
$\beta_0=1400$	Media	1396.447	1396.447	1396.447	1396.446	1396.447
	MSE	1.81587e+03	1.81587e+03	1.81584e+03	1.81587e+03	1.81587e+03
$\beta_1=150$	Media	149.4796	149.4796	149.4795	149.4796	149.4796
	MSE	8.6841e+01	8.6841e+01	8.6842e+01	8.6841e+01	8.6841e+01
$s^2 = 90000$	Media	433.3204	433.3205	430.7399	430.7407	433.3206
	MSE	8.02223e+09	8.02223e+09	8.02269e+09	8.02223e+09	8.02223e+09
$\alpha_1 = 10000$	Media	8039.424	10049.31	8039.556	10049.47	10049.30
	MSE	3.37052e+07	4.66606e+07	3.37042e+07	4.66606e+07	4.66606e+07
$\alpha_2 = 400$	Media	322.5210	403.1517	322.5212	403.1516	403.1516
	MSE	6.64469e+04	9.44527e+04	6.64444e+04	9.44525e+04	9.44529e+04
$\alpha_{12} = 0$	Media	-0.024546	-2.45460e-02	-0.024568	-2.45660e-02	-0.196736
	MSE	2.43011e-01	2.43005e-01	2.43007e-01	2.43005e-01	9.97995e+05

Cuadro 5.8: Simulación de los datos según modelo de Hachemeister con errores MA(1) y $\rho = 0.4$

El cuadros 5.8 muestra los resultados cuando los términos de error siguen una estructura MA(1) con $\rho = 0.4$. Como podemos ver, en este caso, los valores de MSE para el método

ML-MA son menores a los encontrados en con el método clásico, en especial en los valores encontrados en las varianzas de los coeficientes, así también los resultados son similares cuando el método de estimación en el REML. Como podemos ver, en este caso, los valores de MSE para los métodos ML-MA y REML-MA son menores a los encontrados en con el método clásico, en especial en los valores encontrados en las varianzas de los coeficientes. Como se puede observar el método clásico no resulta ser el mejor en la estimación de los parámetros.



Capítulo 6

Conclusiones y Recomendaciones

6.1. Conclusiones

1. Uno de los propósitos de esta tesis fue mostrar que los modelos de credibilidad de Bühlmann (1967), Bühlmann y Straub (1970), Jewell (1975) y Hachemeister (1975) podían ser descritos desde un punto de visto geométrico, con la ayuda de la teoría de espacios de Hilbert y más concretamente con la aplicación del teorema de proyección en estos espacios. Al realizar esto hemos constatado que este enfoque no sólo simplifica de manera significativa la cantidad de cálculos involucrados, sino que también brinda una interpretación más clara y elegante del concepto de prima de Bayes, el cual surge de un enfoque Bayesiano de credibilidad.
2. Un segundo propósito de esta tesis fue evidenciar la relación existente entre los modelos lineales mixtos y la teoría de credibilidad, al mostrar que los programas estadísticos, diseñados especialmente para la estimación de parámetros de los modelos lineales mixtos, podían aplicarse a los datos de credibilidad. Hemos demostrado concretamente que, con la ayuda del lenguaje **R** y el paquete **lme**, la estimación de los parámetros estructurales de credibilidad se puede realizar de una manera rápida y eficiente a través de los métodos ML y REML.
3. En efecto, en términos del MSE, se ha podido observar que la metodología para estimar los parámetros estructurales utilizando los métodos tradicionales no siempre son los más adecuados. Esto se ha hecho más evidente en el caso en que las severidades no eran independientes al interior de las pólizas y poseían una estructura de dependencia, que nosotros asumimos era de medias móviles de primer orden. En estos casos los mejores resultados en términos de sesgo y eficiencia fueron los basados en los modelos lineales mixtos a través del estimador de máxima verosimilitud o máxima verosimilitud restringida.

6.2. Recomendaciones

1. El desarrollo de este trabajo ha abierto la posibilidad de profundizar algunos tópicos que por razones de tiempo y espacio no han podido ser aquí analizados. En primer lugar podríamos hacer un estudio de simulación más exhaustivo fijando otros escenarios tal como un portafolio con una menor cantidad de pólizas pero con una mayor información histórica, la cual podría no estar necesariamente balanceada. Asimismo hemos basado nuestra comparación en base al sesgo y la variabilidad de los estimadores. Un criterio a explorar es también analizar el poder predictivo del modelo para lo cual contamos con la teoría de los BLUP desarrollada en el capítulo 3.
2. Por otro lado sería interesante extender el estudio de simulación a otras estructuras de dependencia más fuertes que la del MA(1). Nosotros intuimos que en estos casos podríamos tener aún mayores diferencias en favor del enfoque de los modelos lineales mixtos.
3. No hemos finalmente abordado todos los diferentes modelos de credibilidad que puedan ser expresados a través de un modelo lineal mixto. Otros modelos tales como el [Dannenburg et al. \(1996\)](#) u otros recientemente introducidos para hacer que estos sean más robustos como los señalados en [Ward \(2008\)](#) podrían ser también analizados.

Apéndice A

Espacios de Hilbert

En este apéndice daremos, en base a [Valdivieso y Lugón \(1999\)](#) una breve presentación de los espacios de Hilbert y del teorema de proyección en estos espacios. Introduciremos también los espacios L^2 y la noción de esperanza condicional.

Una clase muy especial de espacios vectoriales está constituida por los llamados espacios de Hilbert. La ventaja de trabajar en ellos es que al ser estos espacios vectoriales sus elementos se comportan como vectores y por tanto uno puede sacar ventaja de su geometría a fin de obtener de una manera bastante rápida y elegante resultados que algebraicamente podrían resultar complicados.

Definición A.2 *Llamaremos un espacio de producto interno a todo espacio vectorial \mathcal{H} provisto de una aplicación (denominada producto interno):*

$$\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R} \quad \text{tal que}$$

1. $\langle x, x \rangle \geq 0$.
2. $\langle x, x \rangle = 0 \Leftrightarrow x = 0$.
3. $\langle x, y \rangle = \langle y, x \rangle$.
4. $\langle \alpha x + z, y \rangle = \alpha \langle x, y \rangle + \langle z, y \rangle, \quad \forall \alpha \in \mathbb{R}$.

Un espacio \mathcal{H} de producto interno puede verse también como un espacio normado mediante la introducción de la norma:

$$\|x\| = \langle x, x \rangle^{1/2}, \quad \forall x \in \mathcal{H}.$$

Definición A.3 *Una sucesión (x_n) de elementos de un espacio métrico \mathcal{H} converge (en norma) a un elemento $x \in \mathcal{H}$ sii $\|x_n - x\| \rightarrow 0$ conforme $n \rightarrow \infty$.*

Entre las principales propiedades de los espacios en discusión destacan las siguientes.

Proposición A.2 *Sea \mathcal{H} un espacio de producto interno con x, y dos elementos de \mathcal{H} . Entonces:*

- (a) *Desigualdad de Cauchy-Schwartz: $|\langle x, y \rangle| \leq \|x\| \|y\|$.*

(b) *Desigualdad triangular:* $\|x + y\| \leq \|x\| + \|y\|$.

(c) *La ley del paralelogramo:* $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$.

(d) *Continuidad del producto interno:* Si (x_n) e (y_n) son dos sucesiones en \mathcal{H} tales que $\|x_n - x\| \rightarrow 0$ e $\|y_n - y\| \rightarrow 0$ conforme $n \rightarrow \infty$, entonces:

(i) $\|x_n\| \rightarrow \|x\|$ conforme $n \rightarrow \infty$ y

(ii) $\langle x_n, y_n \rangle \rightarrow \langle x, y \rangle$ conforme $n \rightarrow \infty$.

Análogamente al caso euclideo, se puede definir el “ángulo” θ entre dos elementos no nulos x e y de \mathcal{H} mediante:

$$\cos(\theta) = \frac{\langle x, y \rangle}{\|x\| \|y\|}$$

En particular x e y se dirán *ortogonales* si $\langle x, y \rangle = 0$.

Definición A.4 Sea \mathcal{H} un espacio normado. Una sucesión (x_n) en \mathcal{H} es de Cauchy sii $\|x_n - x_m\| \rightarrow 0$ conforme $n, m \rightarrow \infty$. El espacio \mathcal{H} se dice *completo* sii toda sucesión de Cauchy en \mathcal{H} es convergente.

Definición A.5 Un espacio de producto interno \mathcal{H} es de Hilbert sii es completo.

Uno de los resultados más valiosos en la teoría de los espacios de Hilbert lo constituye el **teorema de proyección**. La idea detrás es muy simple. Supongamos que estamos en \mathbb{R}^3 (un resultado estándar establece que \mathbb{R}^p es un espacio de Hilbert, $\forall p \in \mathbb{N}^+$). Dados dos vectores linealmente independientes x_1 y x_2 , podemos considerar el plano generado por ellos:

$$\mathcal{M} = \{\alpha x_1 + \beta x_2 / \alpha, \beta \in \mathbb{R}\}.$$

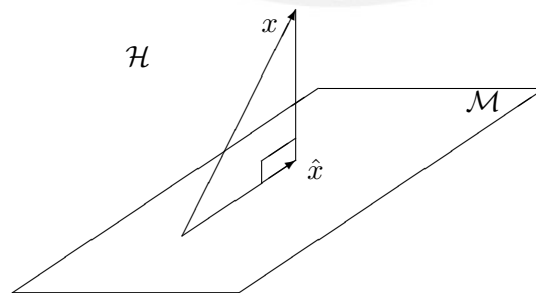


Figura A.1: La proyección de x sobre \mathcal{M}

Si escogemos un elemento arbitrario $x \in \mathbb{R}^3$, estaremos interesados en encontrar un elemento óptimo \hat{x} en \mathcal{M} en el sentido que \hat{x} sea el elemento más cercano en \mathcal{M} a x . Nuestro conocimiento elemental de geometría Euclidea nos dice que \hat{x} viene dado por la proyección ortogonal de x sobre \mathcal{M} (ver la Figura (A.1)). Esto es exactamente lo que el teorema de proyección establece, de trabajarse con el espacio de Hilbert $\mathcal{H} = \mathbb{R}^3$.

Definición A.6 *El complemento ortogonal de un subconjunto \mathcal{M} del espacio de Hilbert \mathcal{H} se define como el conjunto \mathcal{M}^\perp de todos los elementos de \mathcal{H} que sean ortogonales a todo elemento de \mathcal{M} . Esto es*

$$x \in \mathcal{M}^\perp \Leftrightarrow \langle x, y \rangle = 0, \forall y \in \mathcal{M}.$$

Definición A.7 *Un subespacio vectorial \mathcal{M} de un espacio de Hilbert \mathcal{H} es cerrado sii \mathcal{M} contiene a todos sus elementos límites. Vale decir, si para cualquier sucesión $(x_n) \subseteq \mathcal{M}$ tal que $\|x_n - x\| \rightarrow 0$ conforme $n \rightarrow \infty$, se cumple que $x \in \mathcal{M}$.*

Teorema A.7 (Teorema de Proyección) *Si \mathcal{M} es un subespacio cerrado de un espacio de Hilbert \mathcal{H} y $x \in \mathcal{H}$, entonces*

(a) *Existe un único elemento $\hat{x} \in \mathcal{M}$ tal que*

$$\|x - \hat{x}\| = \inf_{y \in \mathcal{M}} \|x - y\|.$$

(b) *$\hat{x} \in \mathcal{M}$ y $\|x - \hat{x}\| = \inf_{y \in \mathcal{M}} \|x - y\| \Leftrightarrow \hat{x} \in \mathcal{M}$ y $x - \hat{x} \in \mathcal{M}^\perp$. En tal sentido, llamaremos a \hat{x} la proyección (ortogonal) de x sobre \mathcal{M} .*

A.1. El espacio L^2

Uno de los más importantes espacios de Hilbert lo constituye el espacio $L^2(E, \mathcal{F}, \mu)$, donde (E, \mathcal{F}, μ) es un espacio medible constituido por un conjunto E , una σ -álgebra \mathcal{F} de subconjuntos de E y una medida μ sobre E . En nuestro trabajo el principal espacio de referencia será el espacio de probabilidad (Ω, \mathcal{F}, P) , en cuyo caso los elementos de $L^2(\Omega, \mathcal{F}, P)$ estarán constituidos por variables aleatorias. Otro ejemplo muy común es cuando E es un subconjunto de \mathbb{R} , \mathcal{F} la σ -álgebra de Borel y μ la medida de Lebesgue en \mathbb{R} .

El espacio $L^2(E, \mathcal{F}, \mu)$ es un espacio de producto interno con

$$L^2(E, \mathcal{F}, \mu) = \{f : \Omega \rightarrow \mathbb{R} \mid f \text{ es } \mathcal{F}\text{-medible y } \int_E f(x)^2 d\mu(x) < \infty\}.$$

y producto interno

$$\langle f, g \rangle = \int_E f(x)g(x) d\mu(x), \quad \forall f, g \in L^2(E, \mathcal{F}, \mu).$$

Es posible probar que $L^2(E, \mathcal{F}, \mu)$ con las operaciones puntuales usuales de multiplicación por un escalar y adición, resulta ser junto a $\langle \cdot, \cdot \rangle$ un espacio de Hilbert. La norma en este espacio viene definida por

$$\|f\| = \sqrt{\int_E f(x)^2 d\mu(x)}.$$

Nótese que si $\omega : E \rightarrow]0, \infty[$ es una función medible tal que $\int_E f(x)^2 \omega(x) d\mu(x) < \infty, \forall f \in L^2(E, \mathcal{F}, \mu)$, entonces el espacio de producto interno

$$\langle f, g \rangle_\omega = \int_E f(x)g(x)\omega(x)d\mu(x)$$

que lo denotaremos por $L^2_\omega(E, \mathcal{F}, \mu)$ es también un espacio de Hilbert, desde que este es idéntico al espacio $L^2(E, \mathcal{F}, \nu)$, donde ν es una nueva medida definida por

$$\nu(A) = \int_A \omega(x) d\mu(x), \quad \forall A \in \mathcal{F}.$$

A este espacio se le suele denominar el espacio $L^2(E, \mathcal{F}, \mu)$ ponderado.

A.2. La esperanza condicional

Consideremos ahora el espacio $L^2(\Omega, \mathcal{F}, P)$ de todas las variables aleatorias con segundo momento finito. Recordemos que este es un espacio de Hilbert con producto interno

$$\langle X, Y \rangle = \int_\Omega X(\omega)Y(\omega)dP(\omega) = E[XY], \quad \forall X, Y \in L^2(\Omega, \mathcal{F}, P)$$

y norma

$$\|X\| = \sqrt{E[X^2]}, \quad \forall X \in L^2(\Omega, \mathcal{F}, P).$$

El valor esperado o esperanza de una variable aleatoria X constituye, por lo usual, un valor representativo de los posibles valores de X . Esta medida no requiere contar con algún grado de información exacta sobre la ocurrencia de los eventos del espacio muestral Ω . De disponerse de información uno entonces podría preguntarse que hacer para encontrar algo similar a $E[X]$; pero que incluya tal información. La respuesta a tal interrogante nos conduce precisamente al concepto de la esperanza condicional de X con respecto a una σ -álgebra \mathcal{G} , donde \mathcal{G} representa la información disponible.

Sea $X \in L^2(\Omega, \mathcal{F}, P)$ y sea \mathcal{G} es una sub- σ -álgebra de \mathcal{F} . Definamos ahora \mathbf{G} como el subespacio de $L^2(\Omega, \mathcal{F}, P)$ constituido por todas las variables aleatorias \mathcal{G} -medibles. No es difícil probar que este es un subespacio cerrado de L^2 . Así, el teorema de proyección garantiza la existencia de una única variable aleatoria $Y \in \mathbf{G}$ tal que

$$\|X - Y\| \leq \|X - Z\|, \quad \forall Z \in \mathbf{G}.$$

Definición A.8 A la variable aleatoria Y anterior se le denota por $E[X|\mathcal{G}]$ y se le llama la esperanza condicional de $X \in L^2$ dada la σ -álgebra \mathcal{G} .

Del teorema de proyección, se desprende también que $E[X|\mathcal{G}]$ es la única variable aleatoria en \mathbf{G} tal que

$$E[(X - E[X|\mathcal{G}])Z] = 0, \quad \forall Z \in \mathbf{G}. \quad (\text{A.1})$$

La esperanza condicional satisface las propiedades siguientes.

Proposición A.3 Sean X e Y dos variables aleatorias y $a \in \mathbb{R}$. Entonces:

- (a) $E[X|\mathcal{F}_0] = E[X]$ (donde $\mathcal{F}_0 = \{\emptyset, \Omega\}$).
- (b) $E[\alpha X + Y|\mathcal{G}] = \alpha E[X|\mathcal{G}] + E[Y|\mathcal{G}]$.
- (c) Si X es \mathcal{G} -medible: $E[XY|\mathcal{G}] = XE[Y|\mathcal{G}]$. En particular, $E[X|\mathcal{G}] = X$.
- (d) Si \mathcal{K} es una sub- σ -álgebra de \mathcal{G} : $E[X|\mathcal{K}] = E[E[X|\mathcal{G}]|\mathcal{K}] = E[E[X|\mathcal{K}]|\mathcal{G}]$.

Aparte de la esperanza condicional otra variable aleatoria de gran utilidad lo constituye la covarianza condicional. Ella está definida como sigue.

Definición A.9 Sean $X, Y \in L^2(\Omega, \mathcal{F}, P)$ y sea \mathcal{G} una sub- σ -álgebra de \mathcal{F} , definimos la covarianza condicional entre X e Y dada la \mathcal{G} mediante

$$\text{Cov}(X, Y|\mathcal{G}) = E[(X - E[X|\mathcal{G}])(Y - E[Y|\mathcal{G}]|\mathcal{G})] = E[XY|\mathcal{G}] - E[X|\mathcal{G}]E[Y|\mathcal{G}].$$

Proposición A.4

$$\text{Cov}(X, Y) = E[\text{Cov}(X, Y|\mathcal{G})] + \text{Cov}(E[X|\mathcal{G}], E[Y|\mathcal{G}]).$$

Demostración: Dado que:

$$\text{Cov}(X, Y|\mathcal{G}) = E[XY|\mathcal{G}] - E[X|\mathcal{G}]E[Y|\mathcal{G}],$$

se tiene que tomando esperanza a ambos lados:

$$\begin{aligned} E[\text{Cov}(X, Y|\mathcal{G})] &= E[E[XY|\mathcal{G}]] - E[E[X|\mathcal{G}]E[Y|\mathcal{G}]] \\ &= E[XY] - E[E[X|\mathcal{G}]E[Y|\mathcal{G}]]. \end{aligned}$$

Por otro lado, como

$$\begin{aligned} \text{Cov}(E[X|\mathcal{G}], E[Y|\mathcal{G}]) &= E[E[X|\mathcal{G}]E[Y|\mathcal{G}]] - E[E[X|\mathcal{G}]]E[E[Y|\mathcal{G}]] \\ &= E[E[X|\mathcal{G}]E[Y|\mathcal{G}]] - E[X]E[Y] \end{aligned}$$

obtenemos de la expresión anterior que:

$$\text{Cov}(X, Y) = E[XY] + E[X]E[Y] = E[\text{Cov}(X, Y|\mathcal{G})] + \text{Cov}(E[X|\mathcal{G}], E[Y|\mathcal{G}])$$

□

Dado que $\text{Var}(X) = \text{Cov}(X, X)$ una consecuencia directa de la proposición anterior es la siguiente.

Corolario A.1

$$\text{Var}(X) = E[\text{Var}(X|\mathcal{G})] + \text{Var}(E[X|\mathcal{G}]).$$

A.3. La esperanza condicional con respecto a un conjunto de variables

Dados n vectores aleatorios X_1, X_2, \dots, X_n podríamos considerar la menor σ -álgebra \mathcal{G} para lo cual todas estos elementos son medibles. Esta se denomina la σ -álgebra generada por X_1, X_2, \dots, X_n . Si $X \in L^2$ (en el sentido que cada una de sus componentes está en L^2) podríamos definir, como lo hemos visto en la subsección anterior la esperanza condicional de un vector X dada la σ -álgebra anterior. Esta suele denotarse denotarse por:

$$E[X|X_1, \dots, X_n] \quad (\text{A.2})$$

y se denomina la esperanza condicional de X dados los vectores aleatorios X_1, X_2, \dots, X_n .

Supongamos que un conjunto de variables aleatorias X_1, X_2, \dots, X_n son continuas o discretas y denotemos por $f_{X_1, X_2, \dots, X_{n-1}, X_n}$ a la función de densidad o probabilidad conjunta de estas variables. En base a esta distribución conjunta, recordemos que se pueden definir distintos elementos como por ejemplo:

- La función de probabilidad o densidad marginal de $(X_1, X_2, \dots, X_{n-1})$:

$$f_{X_1, X_2, \dots, X_{n-1}}(x_1, x_2, \dots, x_{n-1}) = \int f_{X_1, X_2, \dots, X_{n-1}, X_n}(x_1, x_2, \dots, x_{n-1}, x) dx.$$

- La función de probabilidad o densidad condicional de X_n dado que $(X_1, X_2, \dots, X_{n-1}) = (x_1, x_2, \dots, x_{n-1})$:

$$f_{X_n|X_1=x_1, X_2=x_2, \dots, X_{n-1}=x_{n-1}}(x) = \frac{f_{X_1, X_2, \dots, X_{n-1}, X_n}(x_1, x_2, \dots, x_{n-1}, x)}{f_{X_1, X_2, \dots, X_{n-1}}(x_1, x_2, \dots, x_{n-1})}$$

donde la integral arriba se interpreta como una suma en el caso discreto.

En este sentido, y en relación al desarrollo anterior, la esperanza condicional de la variable X_n dadas las variables anteriores, puede escribirse de manera explícita como:

$$E[X_n|X_1 = x_1, X_2 = x_2, \dots, X_{n-1} = x_{n-1}] = \int x f_{X_n|X_1=x_1, X_2=x_2, \dots, X_{n-1}=x_{n-1}}(x)(x) dx.$$

De quedar esta expresión en función de los valores aún no observados del vector aleatorio $(X_1, X_2, \dots, X_{n-1})$, ella se representará por la variable aleatoria $E[X_n|X_1, X_2, \dots, X_{n-1}]$.

Apéndice B

Rutinas en R

En este apéndice incluimos los diferentes programas en R utilizados a lo largo de este trabajo.

B.1. Programa en R para estimar los parámetros con la data de Hachemeister y el modelo de Bühlmann

```
library(actuar)
library(MASS)
require(stats)
library(rpart)
#require(lme4)
library(nlme)
data(hachemeister)
hachemeister
T=12
nc=5
aux1 = matrix(1,1,12)
##### datos de hachemeister #####
datos = matrix(0,5,12)
pesos = matrix(0,5,12)
# matriz de nc=5 contratos y T=12 observaciones por contrato
auxa= matrix(hachemeister,5,25)
for (j in 1:5){
  for (k in 1:12){
    datos[j, k] = auxa[j,k+1]
    pesos[j, k] = auxa[j,k+13] } }
yt =datos
yt2=t(yt)
Y = as.matrix(as.vector(t(yt)))
# ***** estimación lineal para Bühlmann con método ML *****
I = kronecker(matrix(1:nc),t(aux1))
resg = lme(Y~1, random = list(I = pdDiag(~1)), method = 'ML')
aux3 = VarCorr(resg)
mML = fixef(resg)
aML = as.numeric(aux3[1])
s2ML = as.numeric(aux3[2])
ZML = aML*T/(aML*T + s2ML)
parametrosML = matrix(c(mML,aML,s2ML, ZML),1,4)
# ***** estimación lineal para Bühlmann con método REML *****
I = kronecker(matrix(1:nc),t(aux1))
resg2 = lme(Y~1, random = list(I = pdDiag(~1)), method = 'REML')
aux3 = VarCorr(resg2)
mREML = fixef(resg2)
aREML = as.numeric(aux3[1])
s2REML = as.numeric(aux3[2])
ZREML = aREML*T/(aREML*T + s2REML)
parametrosREML = matrix(c(mREML,aREML,s2REML, ZREML),1,4)
# ***** Libreria actuar para Bühlmann *****
fit2=cm(~state, hachemeister, ratios = ratio.1:ratio.12)
parametrosML
```

```

parametrosREML
aux=c(summary(fit2))
aux$cred
aux$unbiased
aux$means

```

Resultados

```

> parametrosML
      [,1]      [,2]      [,3]      [,4]
[1,] 1671.017 57080.68 46040.47 0.9370179
> parametrosREML
      [,1]      [,2]      [,3]      [,4]
[1,] 1671.017 72310.02 46040.47 0.9496143
> aux=c(summary(fit2))
> aux$cred
[[1]]
[1] 0.9496143 0.9496143 0.9496143 0.9496143 0.9496143
> aux$unbiased
portfolio      state
72310.02 46040.47
> aux$means
$portfolio
[1] 1671.017
$state
[1] 2063.833 1510.500 1821.833 1360.333 1598.583

```

B.2. Programa en R para estimar los parámetros con la data de Hachemeister y el modelo de Bühlmann-Straub

```

library(actuar)
library(MASS)
require(stats)
library(rpart)
require(lme4)
library(nlme)
T=12
nc=5
aux1 = matrix(1,1,12)
##### lecturas de datos de hachemeister #####
data(hachemeister)
datos = matrix(0,5,12)
pesos = matrix(0,5,12)
# matriz de nc=5 contratos y T=12 observaciones por contrato
auxa= matrix(hachemeister,5,25)
for (j in 1:5){
  for (k in 1:12){
    datos[j, k] = auxa[j,k+1]
    pesos[j, k] = auxa[j,k+13] } }
yt =datos
yt2=t(yt)
Y = as.matrix(as.vector(t(yt)))
Ws = matrix( t(pesos),12,5) # Matriz de pesos 12x5
SumW=sum(Ws) # W.. es la suma de todos los pesos
sumWc=matrix(colSums(Ws),1,5) # Wj. es la suma de pesos por contrato
W = as.matrix(as.vector(1/Ws))
I = kronecker(matrix(1:nc),t(aux1))
yt2=t(yt)
auxprom2=matrix(colSums(yt2*Ws),1,nc)
meanXjw =auxprom2/sumWc # Xjw : promedio ponderado para cada contrato
Xww=rowSums(meanXjw*sumWc)/SumW # halla Xww
# ***** Libreria actuar para Bühlmann-Straub *****

```



```

fitBS = cm(~state, hachemeister, ratios = ratio.1:ratio.12,
           weights = weight.1:weight.12)
##### estimación lineal para Bühlmann-Straub con método ML ****
data1 = data.frame(Y,I,W)
W=data.frame(W)
vf1 <- varFixed(~ W)
resg <- lme(Y ~ 1, data=data1,random = list(I = pdDiag(~1)),
           method= 'ML',weights=vf1)
aux3 = VarCorr(resg)
mLBSML = fixef(resg)
aLBSML = as.numeric(aux3[1])
s2LBSML = as.numeric(aux3[2])
ZBSML = matrix(0,1,nc)
for (j in 1:nc){
  ZBSML[1,j]=(aLBSML*sumWc[1, j])/ ( aLBSML*sumWc[1, j]+s2LBSML)
}
#ZBSML = aLBSML*T/(aLBSML*T + s2LBSML)
parametrosLBSML = matrix(c(mLBSML,aLBSML,s2LBSML, ZBSML),1,8)
##### estimación lineal para Bühlmann-Straub con método REML ****
resg2 <- lme(Y ~ 1, data=data1,random = list(I = pdDiag(~1)),
           method= 'REML',weights=vf1)
aux4 = VarCorr(resg2)
mLBSREML = fixef(resg2)
aLBSREML = as.numeric(aux4[1])
s2LBSREML = as.numeric(aux4[2])
ZBSREML = matrix(0,1,nc)
for (j in 1:nc){
  ZBSREML[1,j]=(aLBSREML*sumWc[1, j])/ ( aLBSREML*sumWc[1, j]+s2LBSREML)
}
parametrosLBSREML=matrix(c(mLBSREML,aLBSREML,s2LBSREML,ZBSREML),1,8)
# resultados para los 3 metodos en Bühlmann-Straub
parametrosLBSML # estimadores Lineales usando el metodo ML
parametrosLBSREML # estimadores Lineales usando el metodo REML
fitBS$unbiased
fitBS$mean
fitBS$cred

```

Resultados

```

> parametrosLBSML # estimadores Lineales usando el metodo ML
      [,1] [,2] [,3] [,4] [,5] [,6] [,7]
[1,] 1693.433 50214.21 139003664 0.973104 0.8778543 0.8322621 0.5999815
      [,8]
[1,] 0.9287978
> parametrosLBSREML # estimadores Lineales usando el metodo REML
      [,1] [,2] [,3] [,4] [,5] [,6] [,7]
[1,] 1688.756 64859.74 139053560 0.9790427 0.9027215 0.8649837 0.6594756
      [,8]
[1,] 0.9439558
> fitBS$unbiased
      portfolio      state
      89638.73 139120025.93
> fitBS$mean
$portfolio
[1] 1683.713

$state
[1] 2060.921 1511.224 1805.843 1352.976 1599.829

> fitBS$cred
[1] 0.9847404 0.9276352 0.8984754 0.7279092 0.9587911

```

B.3. Programa en R para estimar los parámetros con la data de Hachemeister y el modelo de Jewell

```
library(actuar)
library(MASS)
require(stats)
library(rpart)
require(lme4)
library(nlme)
T=12
nc=5
aux1 = matrix(1,1,12)
##### lecturas de datos de hachemeister #####
data(hachemeister)
datos = matrix(0,5,12)
pesos = matrix(0,5,12)
# matriz de nc=5 contratos y T=12 observaciones por contrato
auxa= matrix(hachemeister,5,25)
for (j in 1:5){
  for (k in 1:12){
    datos[j, k] = auxa[j,k+1]
    pesos[j, k] = auxa[j,k+13] } }
yt =datos
yt2=t(yt)
Y = as.matrix(as.vector(t(yt)))
# ***** Libreria actuar para Jewell *****
X <- cbind(cohort = c(1, 2, 1, 2, 2), hachemeister)
fitJ <- cm(~cohort + cohort:state, data = X, ratios = ratio.1:ratio.12,
  weights = weight.1:weight.12)
Ws = matrix( t(pesos),12,5) # Matriz de pesos 12x5
SumW=sum(Ws) # W.. es la suma de todos los pesos
sumWc = matrix(colSums(Ws),1,5) # Wj. es la suma de pesos por contrato
W = as.matrix(as.vector(1/Ws))
Ia = kronecker(c(1,1,2,2,3),t(aux1))
Ja = kronecker(c(1,2,1,2,2),t(aux1))
yt2=t(yt)
auxprom2=matrix(colSums(yt2*Ws),1,nc)
meanXjw =auxprom2/sumWc # Xjw : promedio ponderado para cada contrato
Xww=rowSums(meanXjw*sumWc)/SumW # halla Xww
# ***** estimación lineal para Jewell con método ML *****
data1 = data.frame(Y,Ja,Ia,W)
W=data.frame(W)
vf1 <- varFixed(~W)
resg3 <- lme(Y~1,data=data1,random=~1|Ja/Ia, method= 'ML', weight=vf1)
aux3 = VarCorr(resg3)
mLJ = fixef(resg3)
aLJ = as.numeric(aux3[2])
bLJ = as.numeric(aux3[4])
s2LJ = as.numeric(aux3[5])
parametrosLJML=matrix(c(mLJ,aLJ,bLJ,s2LJ),1,4)
# ***** estimación lineal para JEWELL con método REML *****
resg4 <-lme(Y~1,data=data1,random=~1|Ja/Ia, method= 'REML', weight=vf1)
aux4 = VarCorr(resg4)
mLJREML = fixef(resg4)
aLJREML = as.numeric(aux4[2])
bLJREML = as.numeric(aux4[4])
s2LJREML = as.numeric(aux4[5])
ZBSREML = matrix(0,1,nc)
parametrosLJREML = matrix(c(mLJREML,aLJREML,bLJREML,s2LJREML),1,4)
# resultados para los 4 metodos en Bühlmann-Straub
parametrosLJML # estimadores Lineales usando el metodo ML
parametrosLJREML # estimadores Lineales usando el metodo REML
fitJ # Libreria actuar
predict(fitJ)
```

Resultados

```
> parametrosLJML # estimadores Lineales usando el metodo ML
      [,1]      [,2]      [,3]      [,4]
[1,] 1742.321 39944.65 12446.72 138587477
> parametrosLJREML # estimadores Lineales usando el metodo REML
      [,1]      [,2]      [,3]      [,4]
[1,] 1793.06 111698.5 0.01000037 150301800
> fitJ          # Libreria actuar
Call:
cm(formula = ~cohort + cohort:state, data = X, ratios = ratio.1:ratio.12,
   weights = weight.1:weight.12)
```

```
Structure Parameters Estimators
Collective premium: 1742.22
```

```
Between cohort variance: 87263.7
Within cohort/Between state variance: 13414.84
Within state variance: 139120026
```

```
> predict(fitJ)
$cohort
[1] 1941.675 1542.765

$state
[1] 2049.733 1522.032 1864.280 1488.504 1587.097
```

B.4. Programa en R para estimar los parámetros con la data de Hachemeister y el modelo de de regresión de Hachemeister

```
library(actuar)
library(MASS)
require(utils)
require(stats)
library(rpart)
require(lme4)
require(nlme)
T=12
nc=5
A = c(0,0,0,0,0,0)
B1 = c(0,0,0,0,0,0)
aux1 = matrix(1,1,12)
##### lecturas de datos de hachemeister #####
data(hachemeister)
datos = matrix(0,5,12)
pesos = matrix(0,5,12)
auxa= matrix(hachemeister,5,25) # matriz de nc=5 contratos y T=12
fith<-cm(~state,hachemeister,regformula=~time,regdata=data.frame(time=1:12),
  ratios = ratio.1:ratio.12, weights = weight.1:weight.12)
#####
# imprime primas
#####
resul=predict(fith, newdata=data.frame(time = 13))
resul # imprime la prima para cada contrato
aa=summary(fith, newdata=data.frame(time = 13))
meB=fith$iterative$state
meA=fith$means$portfolio # imprime parametros de regresion
b0=meA[1]
b1=meA[2]
S2=meB
meC=fith$iterative$portfolio
Varb0=meC[1,1]
Varb1=meC[2,2]
Varb12=meC[2,1]
```

```

A = rbind(A,c(b0,b1,S2,Varb0,Varb1,Varb12 ))
A
##### estimación lineal para Hachemeister con método ML#####
aux1 = matrix(1,1,12)
Y = array(t(hachemeister[,2:13]),c(60,1))
X = array(1:12,c(60,1))
aux1 = matrix(1,1,12)
Sub=kronecker(c(1,2,3,4,5),t(aux1))
W = 1./array(t(hachemeister[,14:25]),c(60,1))
dat = data.frame(Y,X,Sub,W)
vf1 <- varFixed(~W)
resg<-lme(Y~X,data=dat,random=list(Sub=pdSymm(~X)),method='ML',weight=vf1)
aux <- VarCorr(resg)
Varb0ML = as.numeric(aux[1])
Varb0ML
Varb1ML = as.numeric(aux[2])
Varb1ML
S2ML = as.numeric(aux[3])
mLBSML = fixef(resg)
b0ML=mLBSML[1]
b1ML=mLBSML[2]
coef(resg)
B1 = rbind(B1,c(b0ML,b1ML,S2ML,Varb0ML,Varb1ML,Varb12 ))
B1

> resul      # imprime la prima para cada contrato
[1] 2436.752 1650.533 2073.296 1507.070 1759.403

rbind(A,c(b0,b1,S2,Varb0,Varb1,Varb12 ))
> A
  (Intercept)      time
A      0.000  0.00000      0      0.00  0.0000  0.000
  1468.775 32.04892 49870187 24154.17 301.8056 2699.975

resg <- lme(Y~X,data=dat,random=list(Sub=pdSymm(~X)),method='ML',weight=vf1)
> aux <- VarCorr(resg)
> aux
Sub = pdSymm(X)
      Variance      StdDev      Corr
(Intercept) 9.454170e+03  97.23256 (Intr)
X            4.384091e+02  20.93822  1
Residual    4.673394e+07 6836.22240
SML = fixef(resg)
> b0ML=mLBSML[1]
> b1ML=mLBSML[2]
> coef(resg)
  (Intercept)      X
1    1659.649 61.804131
2    1434.511 13.322727
3    1554.929 39.253600
4    1397.374  5.325521
5    1464.114 19.697431
> B1 = rbind(B1,c(b0ML,b1ML,S2ML,Varb0ML,Varb1ML,Varb12 ))
> B1
  (Intercept)      X
B1      0.000  0.00000      0      0.00  0.0000  0.000
  1502.115 27.88068 46733940 9454.17 438.4091 2699.975

```

B.5. Programa para simular los datos para el modelo de Bühlmann

La simulación se realiza en dos escenarios diferentes, el escenario 1 se considera una estructura de errores independientes, luego el escenario 2 considera una estructura de errores MA(1)

```
library(actuar)
library(MASS)
require(stats)
library(rpart)
require(lme4)
library(nlme)
escenario = 1 # escenario 1 = iid , escenario = 2 MA(1)
T = 5 # número de observaciones por póliza
mm = 500 # numero de muestras simuladas
sigma2e = 20 # varianza total de las reclamaciones
m = 400 # prima de riesgo colectiva
Sigma = diag(T)
aux1 = matrix(1,1,5)
A = rep(0,4)
B1 = A
B2 = A
C1 = A
C2 = A
nc = 50 # numero de contratos
for (h in 1:mm){
# Simulación de las severidades
E = mvrnorm(n = nc, rep(0, T), sigma2e*Sigma) # matriz de errores
alpha = mvrnorm(n = nc,0,25) # media para cada contrato
aux2 = kronecker(aux1,alpha)
yt = m + aux2 + E # datos simulados
Y = as.matrix(as.vector(t(yt)))
I = kronecker(matrix(1:nc),t(aux1))
yt2=t(yt)
data1 = data.frame(Y,I)
meanc =matrix(colMeans(yt2),1,50) # promedio por póliza
# ***** estimación clásica de los parámetros de Bulhmann ****
state1=matrix(1:50)
ratio=yt
X <- data.frame(cbind(state1,ratio))
fitB = cm(~X1, data=X, ratios=X2:X6)
mB = as.numeric(fitB$mean[1]) # estima el parámetro m
aB = as.numeric(fitB$unbiased[1])
S2B = as.numeric(fitB$unbiased[2])
ZB =(aB*T)/(S2B+(aB*T)) # factor de credibilidad
# ***** estimación lineal para Bühlmann con método ML *****
resg2 = lme(Y~1,data=data1, random = list(I = pdDiag(~1)), method = 'ML')
aux12 = VarCorr(resg2)
mML = fixef(resg2)
aML = as.numeric(aux12[1])
s2ML = as.numeric(aux12[2])
ZML = aML*T/(aML*T + s2ML)
# ***** estimación lineal para Bühlmann con método REML *****
resg3 = lme(Y~1,data=data1,random=list(I=pdDiag(~1)),method= 'REML')
aux13 = VarCorr(resg3)
mREML = fixef(resg3)
aREML = as.numeric(aux13[1])
s2REML = as.numeric(aux13[2])
ZREML = aREML*T/(aREML*T + s2REML)
# ***** estimación lineal para Bühlmann con método ML-MA(1) *****
eAR <- corARMA(0.4, p = 0, q = 1)
resg4=lme(Y~1,data=data1,random=list(I=pdDiag(~1)),method="ML",corr= eAR)
aux14 = VarCorr(resg4)
```



```

mMLa = fixef(resg4)
aMLa = as.numeric(aux14[1])
s2MLa = as.numeric(aux14[2])
ZMLa = aMLa*T/(aMLa*T + s2MLa)
# ***** estimación lineal para Bühlmann con método REML-MA(1) *****
resg5=lme(Y~1,data=data1,random=list(I=pdDiag(~1)),method="REML",corr=eAR)
aux15 = VarCorr(resg5)
mREMLa = fixef(resg5)
aREMLa = as.numeric(aux15[1])
s2REMLa = as.numeric(aux15[2])
ZREMLa = aREMLa*T/(aREMLa*T + s2REMLa)
# Matrices con valores estimados
A = rbind(A,c(mB, ZB,aB,S2B))
B1 = rbind(B1,c(mML,ZML,aML,s2ML))
B2 = rbind(B2,c(mMLa, ZMLa,aMLa,s2MLa))
C1 = rbind(C1,c(mREML, ZREML,aREML,s2REML))
C2 = rbind(C2,c(mREMLa,ZREMLa,aREMLa,s2REMLa))
} # termino la simulación
# Parámetros
true = matrix(c(400,25/29,25,20),1,5)
# m prima z a s2
sesgoc1 = colMeans(A[-1,]-kronecker(true, matrix(1,mm,1)))
sesgo2 = colMeans(B1[-1,]-kronecker(true, matrix(1,mm,1)))
sesgo2a = colMeans(B2[-1,]-kronecker(true, matrix(1,mm,1)))
sesgo3 = colMeans(C1[-1,]-kronecker(true, matrix(1,mm,1)))
sesgo3a = colMeans(C2[-1,]-kronecker(true, matrix(1,mm,1)))
# cuadrados medios
MSE1 = colMeans( (A[-1,]-kronecker(true, matrix(1,mm,1)))^2)
MSE2 = colMeans( (B1[-1,]-kronecker(true, matrix(1,mm,1)))^2)
MSE2a = colMeans( (B2[-1,]-kronecker(true, matrix(1,mm,1)))^2)
MSE3 = colMeans( (C1[-1,]-kronecker(true, matrix(1,mm,1)))^2)
MSE3a = colMeans( (C2[-1,]-kronecker(true, matrix(1,mm,1)))^2)
par = c("400", " 25/29 ", " 25 ", " 20 ")
ResMEAN= data.frame("PROMEDIOS"= par, "CLÁSICO"= colMeans(A[-1,]),
"LM"=colMeans(B1[-1,]),"LM-MA"=colMeans(B2[-1,]),"REML" = colMeans(C1[-1,]),
"REML-MA" = colMeans(C2[-1,]))
ResMSE = data.frame("MSE" = par, "CLÁSICO" = MSE1, "LM" = MSE2,
"LM-MA" = MSE2a, "REML" = MSE3, "REML-MA" = MSE3a)
ResSESGO = data.frame("SESGO" = par, "CLÁSICO" = sesgoc1, "LM" = sesgo2,
"LM-MA" = sesgo2a, "REML" = sesgo3, "REML-MA" = sesgo3a)
ResMEAN
ResMSE
ResSESGO

```

#Programa para simular por el modelo de Buhlmann y estructura de errores MA(1)

```

library(actuar)
library(MASS)
require(stats)
library(rpart)
require(lme4)
library(nlme)
escenario = 2 # escenario 1 = iid, escenario = 2 MA(1)
T = 5 # número de observaciones por poliza
mm = 500 # numero de muestras simuladas
sigma2e = 20 # varianza total de las reclamaciones
m = 400 # prima de riesgo colectiva
Sigma = diag(T)
aux1 = matrix(1,1,5)
A = rep(0,4)
B1 = A
B2 = A
C1 = A

```



```

C2 = A
nc = 50 # numero de contratos
for (h in 1:mm){
  ma1.model=list(ma=0.4)
  MA2 = c(0,0,0,0,0)
  for (j in 1:nc){
    Sim2=arima.sim(model=ma1.model,n=5,innov=rnorm(5,mean=0,
      sd=sqrt(20/(1+0.4*0.4))))
    MA2 = rbind(MA2,Sim2)
  }
MA1 = matrix(MA2[-1,],nc,5)
# Simulación de las severidades
alpha = mvrnorm(n = nc,0,25)
aux2 = kronecker(aux1,alpha)
yt = m + aux2 + MA1
Y = as.matrix(as.vector(t(yt)))
I = kronecker(matrix(1:nc),t(aux1))
yt2=t(yt)
data1 = data.frame(Y,I)
meanc =matrix(colMeans(yt2),1,50) # promedio por póliza
# ***** estimación clásica de los parámetros de Bulhmann ****
state1=matrix(1:50)
ratio=yt
X <- data.frame(cbind(state1,ratio))
fitB = cm(~X1, data=X, ratios=X2:X6)
mB = as.numeric(fitB$mean[1]) # estima el parámetro m
aB = as.numeric(fitB$unbiased[1])
S2B = as.numeric(fitB$unbiased[2])
ZB =(aB*T)/(S2B+(aB*T)) # factor de credibilidad
# ***** estimación lineal para Bühlmann con método ML *****
resg2=lme(Y~1,data=data1,random=list(I=pdDiag(~1)),method='ML')
aux12 = VarCorr(resg2)
mML = fixef(resg2)
aML = as.numeric(aux12[1])
s2ML = as.numeric(aux12[2])
ZML = aML*T/(aML*T + s2ML)
# ***** estimación lineal para Bühlmann con método REML *****
resg3 = lme(Y~1,data=data1,random=list(I=pdDiag(~1)),method='REML')
aux13 = VarCorr(resg3)
mREML = fixef(resg3)
aREML = as.numeric(aux13[1])
s2REML = as.numeric(aux13[2])
ZREML = aREML*T/(aREML*T + s2REML)
# ***** estimación lineal para Bühlmann con método ML-MA(1) *****
eAR <- corARMA(0.4, p = 0, q = 1)
resg4 =lme(Y~1, data=data1,random=list(I= pdDiag(~1)),method = "ML",
corr = eAR)
aux14 = VarCorr(resg4)
mMLa = fixef(resg4)
aMLa = as.numeric(aux14[1])
s2MLa = as.numeric(aux14[2])
ZMLa = aMLa*T/(aMLa*T + s2MLa)
# ***** estimación lineal para Bühlmann con método REML-MA(1) *****
resg5 =lme(Y~1, data=data1,random=list(I=pdDiag(~1)),method = "REML",
corr = eAR)
aux15 = VarCorr(resg5)
mREMLa = fixef(resg5)
aREMLa = as.numeric(aux15[1])
s2REMLa = as.numeric(aux15[2])
ZREMLa = aREMLa*T/(aREMLa*T + s2REMLa)
# Matrices con valores estimados
A = rbind(A,c(mB, ZB,aB,S2B))
B1 = rbind(B1,c(mML,ZML,aML,s2ML))
B2 = rbind(B2,c(mMLa, ZMLa,aMLa,s2MLa))

```

```

C1 = rbind(C1,c(mREML, ZREML,aREML,s2REML))
C2 = rbind(C2,c(mREMLa,ZREMLa,aREMLa,s2REMLa))
} # termino la simulación
true = matrix(c(400,25/29,25,20),1,5)
# m prima z a s2
sesgoc1 = colMeans(A[-1,]-kronecker(true, matrix(1,mm,1)))
sesgo2 = colMeans(B1[-1,]-kronecker(true, matrix(1,mm,1)))
sesgo2a = colMeans(B2[-1,]-kronecker(true, matrix(1,mm,1)))
sesgo3 = colMeans(C1[-1,]-kronecker(true, matrix(1,mm,1)))
sesgo3a = colMeans(C2[-1,]-kronecker(true, matrix(1,mm,1)))
# cuadrados medios
MSE1 = colMeans( (A[-1,]-kronecker(true, matrix(1,mm,1)))^2)
MSE2 = colMeans( (B1[-1,]-kronecker(true, matrix(1,mm,1)))^2)
MSE2a = colMeans( (B2[-1,]-kronecker(true, matrix(1,mm,1)))^2)
MSE3 = colMeans( (C1[-1,]-kronecker(true, matrix(1,mm,1)))^2)
MSE3a = colMeans( (C2[-1,]-kronecker(true, matrix(1,mm,1)))^2)
par = c("400", " 25/29 ", " 25 ", " 20 ")
ResMEAN = data.frame("PROMEDIOS" = par, "CLÁSICO"= colMeans(A[-1,]),
"LM"=colMeans(B1[-1,]),"LM-MA"=colMeans(B2[-1,]),"REML"=colMeans(C1[-1,]),
"REML-MA" = colMeans(C2[-1,]))
ResMSE = data.frame("MSE" = par, "CLÁSICO" = MSE1, "LM" = MSE2,
"LM-MA" = MSE2a, "REML" = MSE3, "REML-MA" = MSE3a)
ResSESGO = data.frame("SESGO" = par, "CLÁSICO" = sesgoc1, "LM" = sesgo2,
"LM-MA" = sesgo2a, "REML" = sesgo3, "REML-MA" = sesgo3a)
ResMEAN
ResMSE
ResSESGO

```

B.6. Programa para simular los datos para el modelo de Buhlmann-Straub

La simulación de datos en el modelo de Buhlmann-Straub se realiza en dos escenarios diferentes, el escenario 1 considera una estructura de errores independientes, luego el escenario 2 considera una estructura de errores MA(1)

Programa para simular datos, modelo de Buhlmann-Straub con escenarios diferentes

```

BulhmanS <-function(n,k,nsim, m, sigma2e, a, rho,escenario)
{
# Claim simulations under a Bulhman-Straub model with an MA estructura error
# n = num. de reclamos, k = num de pólizas, nsim = número de simulaciones
# m = prima colectiva, sigma2e = total claim variance, a = random effect
# variance
# En el trabajo: Res <-BulhmanS(5,50,500,400,20,25,0.4,2)
# escenario = 2 # escenario 1 = iid, escenario = 2 MA(1)
#####
cf = sqrt(sigma2e/(1 + rho^2))
A = rep(0,5)
B = A
Ba = A
C = A
Ca = A
state1=matrix(1:k)
aux1 = matrix(1,1,n)
I = kronecker(matrix(1:k),t(aux1))
for (j in 1:nsim){
epsilon = rep(0,n)
# uniform simulated weights
W = round(matrix(runif(k*n,5,100),k,n),0) # k x n matriz de pesos simulada
Wx = 1/sqrt(W)
whgi = rowSums(W) # weights sums per policy
ZT = (a*whgi)/(a*whgi + sigma2e) # true credibility factors per policy
if (escenario==1)

```

```

{
  for (i in 1:k){
    eta <- mvrnorm(n,0,20)
    auxepsilon = t(eta)%*%diag(Wx[i,])
    epsilon = rbind(epsilon, auxepsilon)
  }
  epsilon = epsilon[-1,] #errors
  tau = mvrnorm(k,0,a) # heterogeneity parameter
  X = m + kronecker(tau,aux1) + epsilon # severidades simuladas (k x n)
}
if (escenario==2)
{
  for (i in 1:k){
    eta <- mvrnorm(n,0,20)
    auxepsilon = t(eta)%*%diag(Wx[i,])
    epsilon = rbind(epsilon, auxepsilon)
  }
  epsilon = epsilon[-1,] #errors
  tau = mvrnorm(k,0,a) # heterogeneity parameter
  X = m + kronecker(tau,aux1) + epsilon # severidades simuladas (k x n)
}

wm = rowSums(W*X)/whgi # weighted mean
PT=(1-ZT)*m+ZT*wm # primas verdaderas
#Estimación clásica de los parámetros de Bulhmann-Straub ****
Y = data.frame(cbind(state1,X,W))
fitBS = cm(~X1, data=Y, ratios=X2:X6,weights = X7:X11)
mBS = as.numeric(fitBS$mean[1]) # estima el parámetro m
aBS = as.numeric(fitBS$unbiased[1])
S2BS = as.numeric(fitBS$unbiased[2])
ZBS = (aBS*whgi)/(S2BS+(aBS*whgi)) # credibilidad estimadas
PBS=(1-ZBS)*mBS+ZBS*wm # primas estimadas
A = rbind(A, c(mBS, mean(ZBS-ZT), aBS, S2BS, mean(PBS-PT)))
#Estimación por el modelo de Bühlman-Straub con el método ML ****
W0 = as.matrix(as.vector(t(1/W)))
Y = as.matrix(as.vector(t(X)))
data1 = data.frame(Y,I,W0)
vf1 <- varFixed(~W0)
resg2=lme(Y~1,data=data1,random=list(I=pdDiag(~1)),method='ML',weight=vf1)
aux12 = VarCorr(resg2)
mML = fixef(resg2)
aML = as.numeric(aux12[1])
S2ML = as.numeric(aux12[2])
ZML = aML*whgi/(aML*whgi + S2ML)
PML=(1-ZML)*mML+ZML*wm # primas estimadas
B = rbind(B, c(mML, mean(ZML - ZT), aML, S2ML, mean(PML - PT)))
#Estimación lineal para el modelo de Bühlman-Straub con el método ML *****
eMA= corARMA(0.4, p=0, q=1)
resg2a=lme(Y~1,data=data1,random=list(I=pdDiag(~1)),method='ML',weight=vf1,
corr= eMA)
aux22 = VarCorr(resg2a)
mMLa = fixef(resg2a)
aMLa = as.numeric(aux22[1])
S2MLa = as.numeric(aux22[2])
ZMLa = aMLa*whgi/(aMLa*whgi + S2MLa)
PMLa=(1-ZMLa)*mMLa+ZMLa*wm # primas estimadas
Ba = rbind(Ba, c(mMLa, mean(ZMLa - ZT), aMLa, S2MLa, mean(PMLa - PT)))
#Estimación lineal para el modelo de Bühlman-Straub con el método REML ***
resg3 = lme(Y~1,data=data1, random = list(I = pdDiag(~1)),method='REML',
weight=vf1)
aux13 = VarCorr(resg3)
mRML = fixef(resg3)
aRML = as.numeric(aux13[1])
S2RML = as.numeric(aux13[2])
ZRML = aRML*whgi/(aRML*whgi + S2RML)

```

```

PRML=(1-ZRML)*mRML+ZRML*wm # primas estimadas
C = rbind(C, c(mRML, mean(ZRML - ZT), aRML, S2RML, mean(PRML - PT)))
#Estimación lineal para el modelo de Bühlman-Straub con el método REML***
eMA= corARMA(0.4, p=0, q=1)
resg3a=lme(Y~1,data=data1,random=list(I=pdDiag(~1)),method='REML',weight=vf1,
corr=eMA)
aux33 = VarCorr(resg3a)
mRMLa = fixef(resg3a)
aRMLa = as.numeric(aux33[1])
S2RMLa = as.numeric(aux33[2])
ZRMLa = aRMLa*whgi/(aRMLa*whgi + S2RMLa)
PRMLa=(1-ZRMLa)*mRMLa+ZRMLa*wm # primas estimadas
Ca = rbind(Ca, c(mRMLa, mean(ZRMLa - ZT), aRMLa, S2RMLa, mean(PRMLa - PT)))
} # Fin simulaciones
# Parámetros
true = matrix(c(400,0,25.0,20.0, 0),1,5)
par=c("m= 400"," Sesgo Z = 0"," a = 25", "Sigma2 = 20","Sesgo prima = 0")
meanA = colMeans(A[-1,])
MSEA = colMeans( (A[-1,]-kronecker(true, matrix(1,nsim,1)))^2)
meanB = colMeans(B[-1,])
MSEB = colMeans( (B[-1,]-kronecker(true, matrix(1,nsim,1)))^2)
meanBa = colMeans(Ba[-1,])
MSEBa = colMeans( (Ba[-1,]-kronecker(true, matrix(1,nsim,1)))^2)
meanC = colMeans(C[-1,])
MSEC = colMeans( (C[-1,]-kronecker(true, matrix(1,nsim,1)))^2)
meanCa = colMeans(Ca[-1,])
MSECa = colMeans( (Ca[-1,]-kronecker(true, matrix(1,nsim,1)))^2)
Res=data.frame("Parámetro" = par,"Medias_C"=meanA,"MSE_C"=MSEA,"Medias ML"=meanB,
"MSE ML"=MSEB,"Medias ML-MA"=meanBa,"MSE ML-MA"=MSEBa,"Medias REML"=meanC,
"MSE REML"=MSEC,"Medias REML-MA" = meanCa,"MSE REML-MA" = MSECa)
Res}

library(MASS)
library(actuar)
require(stats)
library(nlme)
Res <-BulhmanS(5,50,500,400,20,25,0.4,2)
Res

```

B.7. Programa para simular los datos para el modelo de Jewell

```

#Programa en R para simular datos por el modelo de Jewell y estructura de
errores independientes
JewellSi <-function(n,kp,Po, nsim, m, sigma2e, a, b, rr,escenario)
{
A = rep(0,8)
B = A
Ba = A
C = A
Ca = A
Sigma = matrix(0,5,5)
aux1 = matrix(1,1,n)
aux_1 = matrix(1,1,25)
aux_2 = matrix(1,1,Po)
Eaux = c(0,0,0,0,0)
I = kronecker(matrix(1:k),t(aux1))
Py = c(0,0,0,0,0)
cont=Po*kp
state1=matrix(1:cont)
cohort1=kronecker(matrix(1:Po),t(aux1))
aux01 = matrix(1,5,5)
# simulated weights: Los pesos son generados mediante una Poisson, con lambda
# generado de una Uniforme entre 5 hasta 100
for (i in 1:cont){
lambdaP = runif(1, min=5, max=100)

```

```

P=rpois(5, lambdaP)
pesos = matrix(P,1,n)
Py = rbind(Py,pesos) # Matriz de pesos
}
W = matrix(Py[-1,],cont,n)
Wx = 1/sqrt(W)
whgi = rowSums(W) # weights sums per policy
ZT = (a*whgi)/(a*whgi + sigma2e) # true credibility factors per policy
Zi = b*(rowSums( matrix(ZT,Po,kp)))/(b*(rowSums( matrix(ZT,Po,kp)))+a)
for (j in 1:nsim){
  epsilon = rep(0,n)
  ##### Simulacion de severidades #####
  if (escenario==1)
  {
    for (k2 in 1:cont){
      pesos=Py[k2+1, ]
      aux=(1/pesos)
      for (j2 in 1:5){
        Sigma[j2,j2] = aux[j2]
      }
      Eaux=rbind(Eaux,mvrnorm(n = 1,rep(0,5),16*Sigma))
    }
    Eijt = matrix( Eaux[-1,],cont,n) # Matriz de errores
    ##### simulando los Ei
    alphaEi = mvrnorm(n= Po,0,b)
    Ei = kronecker(aux01,alphaEi[1,])
    for (k2 in 2:Po){
      aux02 = kronecker(aux01,alphaEi[k2,])
    }
    Ei = rbind(Ei,aux02)
    ##### simulando los Eij
    alphaEij = mvrnorm(n=cont,0,a)
    Eij = kronecker(aux1,alphaEij)
    X = m + Ei + Eij + Eijt # Matriz de observaciones de ncx5
  }
  if (escenario==2)
  {
    for (i in 1:cont){
      eta <-arima.sim(list(order=c(0,0,1),ma=rr),n) # innovations
      auxepsilon = eta%%diag(Wx[i,])
      epsilon = rbind(epsilon, auxepsilon)
    }
    epsilon = cf*epsilon[-1,] #errors
    ##### simulando los Ei
    alphaEi = mvrnorm(n= Po,0,b)
    Ei = kronecker(aux01,alphaEi[1,])
    for (k2 in 2:Po){
      aux02 = kronecker(aux01,alphaEi[k2,])
    }
    Ei = rbind(Ei,aux02)
    ##### simulando los Eij
    alphaEij = mvrnorm(n=cont,0,a)
    Eij = kronecker(aux1,alphaEij)
    X = m + Ei + Eij + epsilon # Matriz de observaciones de ncx5
  }
  ##### finalizo la simulacion de severidades #####
  wm =rowSums(W*X)/whgi # weighted mean
  Yizw=rowSums( matrix(ZT*wm,Po,1))/Zi # Promedio para las polizas
  Yzzw=mean(Yizw*Zi)/mean(Zi) # Promedio general
  PT=(1-ZT)*m+ZT*Yizw # primas verdaderas
  PTi=(1-Zi)*m+Zi*Yzzw # primas
#Estimación clásica de los parámetros de Bulhmann-Straub ****
Y = data.frame(cbind(cohort1,state1,X,W))
fitJ = cm(~X1+X1:X2, data=Y, ratios=X3:X7,weights = X8:X12)

```



```

mJ = as.numeric(fitJ$mean[1]) # estima el parámetro m
bJ = as.numeric(fitJ$unbiased[1])
aJ = as.numeric(fitJ$unbiased[2])
S2J = as.numeric(fitJ$unbiased[3])
ZJ = (aJ*whgi)/(S2J+(aJ*whgi)) # credibilidad estimadas
ZiJ = b*(rowSums( matrix(ZJ,Po,kp)))/(b*(rowSums( matrix(ZJ,Po,kp)))+a)
PJ=(1-ZJ)*mJ+ZJ*Yizw # primas estimadas
Pji=(1-ZiJ)*mJ+ZiJ*Yzzw # primas
A = rbind(A, c(mJ, mean(ZJ-ZT), mean(ZiJ-Zi),aJ,bJ,S2J, mean(PJ-PT),
mean(Pji-PTi)))
#Estimación por el modelo de Bühlman-Straub con el método ML (ML-I) ***
Ia = kronecker(matrix(1:Po),t(aux_1))
Ja = kronecker(matrix(1:n),t(aux1))
Ja = matrix(c(Ja,Ja,Ja,Ja,Ja,Ja,Ja,Ja,Ja,Ja),250,1)
W0 = as.matrix(as.vector(t(1/W)))
Y = as.matrix(as.vector(t(X)))
data1 = data.frame(Y,Ia,Ja,W0)
vf1 <- varFixed(~W0)
resg2=lme(Y~1,data=data1,random=~1|Ia/Ja,method= 'ML',weight=vf1)
aux12 = VarCorr(resg2)
mML = fixef(resg2)
aML = as.numeric(aux12[4])
bML = as.numeric(aux12[2])
S2ML = as.numeric(aux12[5])
ZML = aML*whgi/(aML*whgi + S2ML)
ZiML=bML*(rowSums(matrix(ZML,Po,kp)))/(bML*(rowSums(matrix(ZML,Po,kp)))+aML)
PML=(1-ZML)*mML+ZML*Yizw # primas estimadas
PMLi=(1-ZiML)*mML+ZiML*Yzzw # primas
B= rbind(B, c(mML, mean(ZML - ZT), mean(ZiML-Zi),aML,bML,S2ML,mean(PML-PT),
mean(PMLi-PTi)))
#Estimación lineal para el modelo de Jewell con el método ML (ML-MA) *****
eMA= corARMA(0.4, p=0, q=1)
resg2a=lme(Y~1,data=data1,random=~1|Ia/Ja,method='ML',weight=vf1,corr= eMA)
aux2a = VarCorr(resg2a)
mMLa = fixef(resg2a)
aMLa = as.numeric(aux2a[4])
bMLa = as.numeric(aux2a[2])
S2MLa = as.numeric(aux2a[5])
ZMLa = aMLa*whgi/(aMLa*whgi + S2MLa)
ZiMLa = bMLa*(rowSums( matrix(ZMLa,Po,kp)))/(bMLa*(rowSums( matrix(ZMLa,Po
,kp)))+aMLa)
PMLa=(1-ZMLa)*mMLa+ZMLa*Yizw # primas estimadas
PMLia=(1-ZiMLa)*mMLa+ZiMLa*Yzzw # primas
Ba = rbind(Ba, c(mMLa, mean(ZMLa - ZT), mean(ZiMLa-Zi), aMLa,bMLa, S2MLa,
mean(PMLa - PT),mean(PMLia-PTi)))
#Estimación lineal para el modelo de Jewell con el método REML (REML-I) *****
vf1 <- varFixed(~W0)
resg3 = lme(Y~1,data=data1, random= ~1|Ia/Ja, method= 'REML',weight=vf1)
aux13 = VarCorr(resg3)
mREML = fixef(resg3)
aREML = as.numeric(aux13[4])
bREML = as.numeric(aux13[2])
S2REML = as.numeric(aux13[5])
ZREML = aREML*whgi/(aREML*whgi + S2REML)
ZiREML = bREML*(rowSums( matrix(ZREML,Po,kp)))/(bREML*(rowSums( matrix(ZREML,
Po,kp)))+aREML)
PREML=(1-ZREML)*mREML+ZREML*Yizw # primas estimadas
PREMLi=(1-ZiREML)*mREML+ZiREML*Yzzw # primas
C = rbind(C, c(mREML, mean(ZREML-ZT), mean(ZiREML-Zi), aREML,bREML, S2REML,
mean(PREML - PT),mean(PREMLi-PTi)))
#Estimación lineal para el modelo de JEWELL con el método REML (REML-MA) ****
eMA= corARMA(0.4, p=0, q=1)
resg3a= lme(Y~1,data=data1,random=~1|Ia/Ja,method='REML',weight=vf1,corr=eMA)
aux13a = VarCorr(resg3a)

```



```

mREMLa = fixef(resg3a)
aREMLa = as.numeric(aux13a[4])
bREMLa = as.numeric(aux13a[2])
S2REMLa = as.numeric(aux13a[5])
ZREMLa = aREMLa*whgi/(aREMLa*whgi + S2REMLa)
ZiREMLa=bREMLa*(rowSums(matrix(ZREMLa,Po,kp)))/(bREMLa*(rowSums(matrix(ZREMLa,
Po,kp)))+aREMLa)
PREMLa=(1-ZREMLa)*mREMLa+ZREMLa*Yizw # primas estimadas
PREMLia=(1-ZiREMLa)*mREMLa+ZiREMLa*Yzzw # primas
Ca=rbind(Ca,c(mREMLa,mean(ZREMLa - ZT),mean(ZiREMLa-Zi),aREMLa,bREMLa,S2REMLa,
mean(PREMLa - PT),mean(PREMLia-PTi)))
} # Fin simulaciones

# Parámetros
true = matrix(c(30,0,0,1,25,16,0 , 0),1,8)
par = c("m = 30"," Sesgo Z = 0"," Sesgo Z =0" , " a = 1"," b = 25","Sigma2=16",
"Sesgo prima = 0", "Sesgo prima = 0")
meanA = colMeans(A[-1,])
MSEA = colMeans( (A[-1,]-kronecker(true, matrix(1,nsim,1)))^2)
meanB = colMeans(B[-1,])
MSEB = colMeans( (B[-1,]-kronecker(true, matrix(1,nsim,1)))^2)
meanBa = colMeans(Ba[-1,])
MSEBa = colMeans( (Ba[-1,]-kronecker(true, matrix(1,nsim,1)))^2)
meanC = colMeans(C[-1,])
MSEC = colMeans( (C[-1,]-kronecker(true, matrix(1,nsim,1)))^2)
meanCa = colMeans(Ca[-1,])
MSECa = colMeans( (Ca[-1,]-kronecker(true, matrix(1,nsim,1)))^2)
Res=data.frame("Parámetro"=par,"Medias_C"=meanA,"MSE_C"=MSEA,"Medias ML"=meanB,
"MSE ML"=MSEB,"Medias ML-MA"=meanBa,"MSE ML-MA"=MSEBa,"Medias REML"=meanC,
"MSE REML"=MSEC,"Medias REML-MA" = meanCa,"MSE REML-MA" = MSECa)
Res}

library(MASS)
library(actuar)
require(stats)
library(nlme)
Res <-JewellSi(5,5,50,500,30,16,1,25,0.4,2)
Res

```

B.8. Programa en R para simular los datos con el modelo de regresión de Hachemeister

#Programa en R para simular datos por el modelo de Regresión de Hachemeister en dos escenarios diferentes: escenario = 1 para estructura de errores independientes y escenario = 2 para una estructura MA(1)

```

library(actuar)
library(MASS)
require(utils)
require(stats)
library(rpart)
require(lme4)
require(nlme)
data(hachemeister)
for (ill in 1:12){
}
escenario = 1 # escenario 1 = iid , escenario = 2 MA(1)
rr=0.4
rho=matrix(rr,12,12)
for (ill in 1:12){
rho[ill,ill]=1
}
RA = rep(0,6)
RB1 = RA

```

```

RB2 = RA
RB3 = RA
RB4 = RA
aux1 = matrix(1,1,12)
Sigma = matrix(0,12,12)
b0R=1400
b1R=150
S2R=90000
Varb0R=10000
Varb1R=400
Varb12R=0
Eaux = c(0)
mm=500
n=12
cf = sqrt(S2R/(1 + rr^2))
for (kno in 1:mm){
  Y = matrix(0,60,1)
  # simula los coeficientes de BETA
  F=matrix(c(Varb0R,Varb12R,Varb12R,Varb1R),2,2)
  beta=matrix(c(b0R,b1R),2,1)
  betac=mvrnorm(n = 5, beta, F)
  # termino la simulacion de los BETA
  aux1 = matrix(1,60,1)
  X1 = array(1:12,c(60,1))
  X=matrix(c(aux1,X1),60,2)
  cont=1
  for (i in 1:5) {
    for (j in 1:12) {
      ab0=betac[i,1]
      ab1=betac[i,2]
      Y[cont]=ab0+ab1*j
      cont=cont+1
    }
  }
  ##### SIMULACION DE LOS DATOS #####
  if (escenario==1)
  {
    Eaux=matrix(0,1,12)
    for (k2 in 1:5){
      pesos=array(t(hachemeister[k2,14:25]),12,1)
      aux=(1/pesos)
      for (j in 1:12){
        Sigma[j, j] = aux[j]
      }
      Eaux=rbind(Eaux,mvrnorm(n = 1,rep(0,12),90000*Sigma))
    }
    Eit = matrix( Eaux[-1,],60,1)
    Y = Y + Eit
  }
  if (escenario==2)
  {
    epsilon = rep(0,n)
    # generando los errores
    for (k2 in 1:5){
      pesos=array(t(hachemeister[k2,14:25]),12,1)
      aux=(1/pesos)
      for (j in 1:12){
        Sigma[j, j] = aux[j]
      }
    }
    # Eaux=rbind(Eaux,mvrnorm(n = 1,rep(0,12),90000*Sigma))
    eta <-arima.sim(list(order=c(0,0,1),ma=rr),n) # innovations
    auxepsilon = eta%%Sigma
    epsilon = rbind(epsilon, auxepsilon)
  }
}

```

```

    epsilon = cf*epsilon[-1,] #errors
#####      Eit = matrix( Eaux[-1,],60,1)
Y = Y + matrix(epsilon,60,1)
}
##### FIN DE LA SIMULACION #####3333
auxy=t(matrix(Y,12,5))
for (kf in 1:5){
  for (kc in 2:13){
    hachemeister[kf,kc]=auxy[kf,kc-1]
  }
}
fith<-cm(~state, hachemeister, regformula=~time, regdata=data.frame(time=1:12),
ratios = ratio.1:ratio.12, weights = weight.1:weight.12)
meB=fith$iterative$state
meA=fith$means$portfolio # imprime parametros de regresion
b0=meA[1]
b1=meA[2]
S2=meB
meC=fith$iterative$portfolio
Varb0=meC[1,1]
Varb1=meC[2,2]
Varb12=meC[2,1]
RA = rbind(RA,c(b0,b1,S2,Varb0,Varb1,Varb12))
##### estimación lineal para Hachemeister con método ML #####
auxL1= matrix(1,1,12)
X = array(1:12,c(60,1))
Sub=kronecker(c(1,2,3,4,5),t(auxL1))
W = 1/array(t(hachemeister[,14:25]),c(60,1)) # inversa de pesos
dat = data.frame(Y,X,Sub,W)
vf1 <- varFixed(~W)
resg <- lme(Y~ X,data=dat,random=list(Sub=pdSymm(~ X)),method="ML",weight=vf1)
RESHL <- VarCorr(resg)
Varb0ML = as.numeric(RESHL[1])
Varb1ML = as.numeric(RESHL[2])
Varb12ML = as.numeric(RESHL[2,3])
S2ML = as.numeric(RESHL[3])
mLBSML = fixef(resg)
b0ML=mLBSML[1]
b1ML=mLBSML[2]
coef(resg)
RB1=rbind(RB1,c(b0ML,b1ML,S2ML,Varb0ML,Varb1ML,Varb12ML))
resg2<-lme(Y~X,data=dat,random=list(Sub=pdSymm(~X)),method='REML',weight=vf1)
aux2 <- VarCorr(resg2)
Varb0REML = as.numeric(aux2[1])
Varb1REML = as.numeric(aux2[2])
Varb12REML = as.numeric(aux2[2,3])
S2REML = as.numeric(aux2[3])
mLBSREML = fixef(resg2)
b0REML=mLBSREML[1]
b1REML=mLBSREML[2]
coef(resg2)
RB2=rbind(RB2,c(b0REML,b1REML,S2REML,Varb0REML,Varb1REML,Varb12REML))
# ***** estimación lineal para Hachemeister con método ML-MA **
eAR <- corARMA(0.4, p = 0, q = 1)
resg3 <- lme(Y~X,data=dat,random=list(Sub=pdSymm(~X)),method='ML',
weight=vf1,corr = eAR)
aux3 <- VarCorr(resg3)
Varb0MLa = as.numeric(aux3[1])
Varb1MLa = as.numeric(aux3[2])
Varb12MLa = as.numeric(aux3[2,3])
S2MLa = as.numeric(aux3[3])
mLBSMLa = fixef(resg3)
b0MLa=mLBSMLa[1]
b1MLa=mLBSMLa[2]

```

```

coef(resg3)
RB3 = rbind(RB3,c(bOMLa,b1MLa,S2MLa,VarbOMLa,Varb1MLa,Varb12MLa ))
# ***** estimación lineal para Hachemeister con método REML-MA **
resg4 <- lme(Y~X,data=dat,random=list(Sub=pdSymm(~X)),
method='REML',weight=vf1,corr = eAR)
aux4 <- VarCorr(resg4)
Varb0REMLa = as.numeric(aux4[1])
Varb1REMLa = as.numeric(aux4[2])
Varb12REMLa = as.numeric(aux4[2,3])
S2REMLa = as.numeric(aux4[3])
mLBSREMLa = fixef(resg4)
b0REMLa=mLBSREMLa[1]
b1REMLa=mLBSREMLa[2]
coef(resg4)
RB4=rbind(RB4,c(b0REMLa,b1REMLa,S2REMLa,Varb0REMLa,Varb1REMLa,Varb12REMLa))
}
true = matrix(c(1400,150,90000,10000,400,0),1,6)
# cuadrados medios
MSE1 = colMeans( (RA[-1,]-kronecker(true, matrix(1,mm,1)))^2)
MSE2 = colMeans( (RB1[-1,]-kronecker(true, matrix(1,mm,1)))^2)
MSE3 = colMeans( (RB2[-1,]-kronecker(true, matrix(1,mm,1)))^2)
MSE4 = colMeans( (RB3[-1,]-kronecker(true, matrix(1,mm,1)))^2)
MSE5 = colMeans( (RB4[-1,]-kronecker(true, matrix(1,mm,1)))^2)
meanA = colMeans(RA[-1,])
meanB1 = colMeans(RB1[-1,])
meanB2 = colMeans(RB2[-1,])
meanB3 = colMeans(RB3[-1,])
meanB4 = colMeans(RB4[-1,])
par = c("1400", " 150 " , " 90000 " , " 10000 " , " 400 " , " 0 ")
ResMEAN = data.frame("PROMEDIOS" = par, "CLÁSICO"= meanA, "LM" = meanB1,
"LM-MA" = meanB2, "REML" = meanC1, "REML-MA" = meanC2)
ResMSE = data.frame("MSE" = par, "CLÁSICO" = MSE1, "LM" = MSE2,
"LM-MA" = MSE2a, "REML" = MSE3, "REML-MA" = MSE3a)
ResMEAN
ResMSE

```

Bibliografía

- Bailey, A. L. (1945). *A generalized theory of credibility*, Proceedings of the casualty actuarial society, vol 13, p. 13-20.
- Boland, P. J. (2007). *Statistical and probabilistic methods in actuarial science*, Chapman and Hall/CRC.
- Bühlmann, H. (1967). *Experience rating and credibility*, ASTIN Bulletin.
- Bühlmann, H. y Gisler, A. (n.d.). *A Course in credibility theory and its applications*, Springer.
- Bühlmann, H. y Straub, E. (1970). *Experience rating and credibility*, Mitteilungen der vereinigung schweizerischer versicherungsmathematiker.
- Dannenburg, D., Kaas, R. y Goovaerts, M. (1996). *Practical Actuarial Credibility Models*, Institute of actuarial science and econometrics, University of Amsterdam, Amsterdam, the Netherlands.
- Dubey, A., G. A. (1981). *On Parameter Estimators in Credibility*, MVSVM, 81:2, 187-212.
- Faraway, J. J. (2006). *Extending the linear model with R*, Chapman and Hall/CRC.
- Frees, E., Young, V. y Luo, Y. (1999). A longitudinal data analysis interpretation of credibility models, *Insurance Mathematics and Economics* **24**: 229–247.
- Fung, W. y Xu, X. (2010). *Estimation and Robustness of Linear Mixed Models in Credibility Context*, Variance: Journal of the casualty actuarial society, Vol 2, p. 66-80.
- Garrido, J. y Zhou, J. (2006). *Credibility theory for generalized linear and mixed models*, NSERC.
- Goovaerts, M. y Hoogstad, W. (1987). *Credibility theory, surveys of actuarial studies*, Nationale-Nederlanden.
- Goovaerts, M., Kaas, R., Vaan Heerwaarden, A. y Bauwelinck, T. (1990). *Effective actuarial methods (Insurance Series, No 3)*, Elsevier Publisher B.V.
- Hachemeister, C. (1975). Credibility for regression models with application to trend.
- Jewell, W. (1975). *The use of collateral data in credibility theory: A hierarchical model*, Giornale dell Istituto degli Attuari Italiani 38,1-16.
- Jiang, J. (2006). *Linear and generalized linear mixed models and their applications*, Springer.
- Katrien, A. (2007). Statistical tools for non-life insurance: Essays on claims reserving and ratemaking for panels and fleets.
- Katrien, A. y Beirlant, J. (2006). Actuarial statistical with generalized linear mixed models, *Insurance Mathematics and Economics*.

- Laird, N. y Ware, J. (1982). *Random-effects models for longitudinal data*, International Biometric Society.
- Manterola, L. (2007). *Métodos de Estadística Espacial para evaluar la influencia de factores medioambientales sobre la incidencia y mortalidad por cáncer*, Universidad de Girona.
- Mayerson, A. (1964). *A Bayesian view of credibility*, Proceedings of the casualty actuarial society, vol 51, p. 85-104.
- Mikosch, T. (2003). *Non-Life insurance mathematics a primer*, University of Copenhagen.
- Moreno, M. T. y Ramos, B. (2003). *Aplicación de los Modelos de credibilidad para el cálculo de primas en el seguro de automóviles*, Comisión Nacional de Seguros y Fianzas.
- Nelder, J. A. y Verrall, R. J. (1997). Credibility theory and generalizad linear models, *Austin Bulletin* 27: 71–82.
- Ohlsson, E. (2006). *Credibility estimators in multiplicative models*, Stockholm University.
- Ordoñez, E. N. (2010). *Estimación de intervalos de confianza para reservas no reportadas en compañías aseguradoras*, PUCP.
- Palma, J. (2006). *Metodología para el diseño de un laboratorio actuarial soportado por tecnología de la información y administración del riesgo*, Master's thesis, Universidad de los Andes.
- Pinheiro, J. y Bates, D. (2000). *Nonlinear mixed-effects models; basic concepts and motivating examples*, Springer, 234.
- Pons, A. (1991). *La teoría de credibilidad y su aplicación a los seguros colectivos*, Master's thesis, Universidad de Barcelona.
- Sarabia, J., Gómez, E. y Vásquez, F. (2006). *Estadística actuarial, teoría y aplicaciones*, vol 37, p. 7-23, Pearson.
- Valdivieso, L. H. (2011). *Notas de clase del curso de Análisis de riesgos*, Maestria en estadística, Sección Matemáticas PUCP.
- Valdivieso, L. y Lugón, A. (1999). *La esperanza condicional en espacios de probabilidad finitos*, Reporte de investigación No 1. Sección Matemáticas PUCP.
- Ward, C. (2008). *Robust theory applied to jewell's hierarchical credibility mode1*, Master's thesis, Concordia University.
- Xu, X. (2008). *Estimation of structural parameters in credibility context using mixed effects models*, Master's thesis, The University of Hong Kong.